

# «Fraktale Analyse von Mammographien»

Diplomabschlussarbeit  
Institut für Medien- und Phototechnik  
Fachhochschule Köln

Autor: Kai Elbers  
Mat.-Nr.: 11045661

Referent: Prof. Dr. Christian Blendl  
Koreferent: Prof. Dr. Gregor Fischer

Köln, August 2008

# «Fractal Analysis of Mammographies»

Diploma Thesis  
Institute of Media and Photo Technology  
University of Applied Sciences Cologne

Author: Kai Elbers  
Mat.-No.: 11045661

First Reviewer: Prof. Dr. Christian Blendl  
Second Reviewer: Prof. Dr. Gregor Fischer

Cologne, August 2008

## **Kurzbeschreibung**

Titel:	Fraktale Analyse von Mammographien
Autor:	Kai Elbers
Referenten:	Prof. Dr. Christian Blendl und Prof. Dr. Gregor Fischer
Zusammenfassung:	<p>Ziel dieser Arbeit ist es, herauszufinden, ob der Parameter der fraktalen Dimension in Zukunft nützlich sein könnte für die computergestützte Auswertung von Mammographien. Dies gilt insbesondere für die Detektion architektonischer Störungen im Brustparenchym, die zuweilen durch den Radiologen schwer ausfindig zu machen sind.</p> <p>Die fraktale Dimension ist deshalb als Parameter so interessant, weil sie auf der Eigenschaft der Selbstähnlichkeit beruht. Eine Selbstähnlichkeit in der Struktur könnte auch das Brustgewebe haben, da Gewebewachstum ein fortlaufender Prozess ist, bei dem sich wiederholend die gleichen Formen herausbilden. Kommt es hier zu pathologischen Veränderungen, so ist dieser Ablauf unterbrochen und die Strukturen verändern sich.</p>
Stichwörter	Mammographie, Fraktale Dimension, Computer Aided Detection
Sperrvermerk	Die vorgelegte Arbeit unterliegt keinem Sperrvermerk.
Datum	28.08.2008

## **Abstract**

Title:	Fractal Analysis of Mammographies
Author:	Kai Elbers
Reviewers:	Prof. Dr. Christian Blendl and Prof. Dr. Gregor Fischer
Summary:	<p>Goal of this thesis is to find out if the fractal dimension could be a useful option in Computer Aided Detection of mammographies. Especially for detection of architectural distortion in the breast parenchyma, which is hard to find, even for an experienced radiologist.</p> <p>The fractal dimension is therefore an interesting approach, since it is based on the quality of selfsimilarity. The breast parenchyma could be selfsimilar as well, because the tissue growth is a continuous process, creating similar structures. If this process is interrupted due to cancer, there could be a change in the structure, which possibly could be detected by the parameter of fractal dimension.</p>
Keywords	Mammography, Fractal Dimension, Computer Aided Detection
Closing Remark	The presented thesis is not subject to a blocking period.
Date	28.08.2008

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>6</b>
1.1	Bildgebende Verfahren in der medizinischen Diagnostik	7
1.1.1	Projektionsradiographie . . . . .	7
1.1.2	Computer-Tomographie . . . . .	9
1.1.3	Magnet-Resonanz-Tomographie und Sonographie	9
1.2	Die Anatomie der weiblichen Brust . . . . .	11
1.3	Das Mammakarzinom . . . . .	11
1.3.1	Epidemiologie . . . . .	13
1.3.2	Symptomatik . . . . .	13
1.3.3	Einstufungen des Mammakarzinoms . . . . .	15
1.4	Die Mammographie . . . . .	15
<b>2</b>	<b>Die fraktale Dimension</b>	<b>19</b>
2.1	Das Küstenlinien-Problem . . . . .	20
2.2	Der Richardson-Effekt . . . . .	21
2.3	Fraktale Dimension von Oberflächen . . . . .	21
2.4	Fraktale in der Natur . . . . .	23
2.5	Schätzung der fraktalen Dimension . . . . .	24
2.5.1	Methode nach Caldwell . . . . .	24
2.5.2	Circular Average Power Spectrum Method . . .	25
2.5.3	Triangular Prism Method . . . . .	25
2.5.4	Differential Box-Counting Method . . . . .	28
2.5.5	Variogram Method . . . . .	29
2.5.6	Isarithm Method . . . . .	30
2.5.7	Methodenkritik . . . . .	30
2.5.8	Vergleich der Caldwell und CAPS Methode . . .	34
2.5.9	Vergleich weiterer Methoden . . . . .	36
<b>3</b>	<b>Material und Methoden</b>	<b>42</b>

---

3.1	Digital Database for Screening Mammography . . . . .	42
3.1.1	Scanner . . . . .	44
3.1.2	Originaldaten . . . . .	45
3.1.3	Konvertierung der Mammographien . . . . .	45
3.2	Matlab . . . . .	45
3.2.1	Einleitung . . . . .	45
3.2.2	Funktionen . . . . .	47
3.2.3	Graphical User Interface . . . . .	47
3.3	Programm CompareFractalEstimation.m . . . . .	47
3.4	Programm MammoFractalAnalyzer.m . . . . .	50
<b>4</b>	<b>Messung und Ergebnisse</b>	<b>53</b>
4.1	Messkriterien . . . . .	53
4.2	Messung . . . . .	54
4.3	Ergebnisse . . . . .	55
4.3.1	Messung A . . . . .	56
4.3.2	Messung B . . . . .	58
4.3.3	Messung C . . . . .	60
4.4	Statistische Analyse . . . . .	62
4.4.1	t-Test . . . . .	62
4.4.2	Ergebnisse t-Test . . . . .	63
<b>5</b>	<b>Diskussion und Fehleranalyse</b>	<b>66</b>
	<b>Literaturverzeichnis</b>	<b>68</b>
	<b>Abbildungsverzeichnis</b>	<b>71</b>
	<b>Tabellenverzeichnis</b>	<b>72</b>
	<b>Eidesstattliche Erklärung</b>	<b>73</b>

# 1 Einführung

Computer Aided Detection [CAD] ist ein in den letzten Jahren zunehmend wichtiges Thema in der Medizin geworden. Hierbei werden mittels Software digitalisierte Aufnahmen verschiedener bildgebender Verfahren eines Patienten untersucht und suspekte Bereiche im Bild gekennzeichnet. Die Entwicklung immer präziser arbeitender Algorithmen macht sich bei der Diagnose verschiedener Krankheiten deutlich bemerkbar und steigert die Sensitivität der Detektion [15]. Speziell bei der Mammographie werden solche Systeme eingesetzt, um den Radiologen bei der Auswertung der Röntgenaufnahmen zu unterstützen bzw. um als „second reader“ zu fungieren.

Immer noch werden viele Mammakarzinome im Anfangsstadium nicht erkannt. Dies kann mehrere Gründe haben, wie zum Beispiel die Überlagerung der Störung durch das Brustparenchym<sup>1</sup> in der Abbildung, technische Fehler oder eine schlechte Aufnahmequalität. Ein weiterer und nicht unwichtiger Grund für die Nichterkennung einer Läsion<sup>2</sup> kann aber auch die schlichte Unaufmerksamkeit oder Unerfahrenheit des Radiologen sein.

Dies alles sind Gründe, die technischen Systeme zur Erkennung von Brustkrebs stetig zu verbessern, die auffälligen Regionen in Mammographien zu kennzeichnen und vielleicht sogar eine Unterscheidung zwischen benigne<sup>3</sup> und maligne<sup>4</sup> Tumoren durch entsprechende Software zu ermöglichen.

CAD wird schon seit längerer Zeit in der Medizin erfolgreich eingesetzt und die Effizienz des Einsatzes wurde in verschiedenen Studien nach-

---

<sup>1</sup>Drüsengewebe

<sup>2</sup>Schädigung anatomischer Struktur

<sup>3</sup>Gutartige

<sup>4</sup>Bösartige

## 1 Einführung

---

gewiesen. So gab es zum Beispiel in Großbritannien eine Studie, in der 10267 Mammogramme untersucht worden sind, die von Frauen ab 50 zwischen 1996 und 1999 erstellt wurden [15]. Die Aufnahmen wurden digitalisiert und von acht erfahrenen Radiologen mithilfe eines CAD Systems erneut untersucht. Die Begutachtung mit CAD ergab eine 6,5 % höhere Detektion der Karzinome als bei der Doppelbegutachtung, bei der die Meinung eines zweiten Radiologen eingeholt wird.

Inzwischen haben sich auch kommerzielle Anbieter von CAD Software auf dem Markt etabliert. Anbieter wie R2 Technologie und ICad entwickelten Systeme, die heute schon in der Medizin eingesetzt werden und aus medizinischer und ökonomischer Sicht sinnvoll erscheinen.

## 1.1 Bildgebende Verfahren in der medizinischen Diagnostik

### 1.1.1 Projektionsradiographie

Es gibt verschiedene radiologische bildgebende Verfahren, die in der Medizin verwendet werden. Die älteste und auch bekannteste ist wohl die Projektionsradiographie, bei der der Körper des Patienten mithilfe hochenergetischer Röntgenstrahlung durchstrahlt wird. Die verschiedenen Organe im Körperbau absorbieren diese Strahlung in unterschiedlichem Maße, wodurch Informationen über den Aufbau des Körpers und der inneren Organe gewonnen werden. Je nachdem was untersucht werden soll, muss man mit „weicher“ oder „harter“ Strahlung arbeiten. Weiche Strahlung liegt im Energiebereich bis 100 KeV [Kiloelektronenvolt], harte Strahlung zwischen 100 und 1000 KeV und ultraharte Strahlung bei über 1 MeV. Bei einem Gewebe wie zum Beispiel bei der weiblichen Brust wird mit weicher Strahlung gearbeitet. In der Röntgendiagnostik spricht man bis 35 kV Röntgenröhrenspannung von Weichstrahltechnik und ab 120 kV von Hartstrahltechnik. Bei der Mammographie wird die Weichstrahltechnik verwendet. Dadurch werden Kontraste innerhalb des Gewebes deutlich sichtbar dar-



## 1 Einführung

---

gestellt. Allerdings wird auch ein Großteil der Strahlung vom Gewebe absorbiert, wodurch eventuell nachhaltige Schäden auftreten können. Die Absorption wird als Energiedosis (Gleichung 1.1) bezeichnet, die in Gray [Gy] gemessen wird und Informationen über die Strahlenbelastung hergibt.

Die Strahlungsqualität ist zusätzlich zur eingestellten Spannung auch abhängig von der Filterung. Eine Filterung wird dann eingesetzt, wenn eine Aufnahme von inhomogenen Körperregionen gemacht wird. Zum Beispiel werden Filter bei Aufnahmen des Thorax<sup>5</sup> verwendet, wo ansonsten ein zu hoher Kontrast im Bild entstünde, da die Lungenflügel die Strahlung kaum absorbieren, die Rippen jedoch sehr stark.

Die Energiefluenz  $\Psi$  der Röntgenstrahlung, die auch als Röntgenstrahlungsintensität bezeichnet wird, kann durch die Gleichung 1.2 ermittelt werden. Sie nimmt mit zunehmender Schichtdicke  $d$  des Körpers und hohem Absorptionskoeffizienten  $\mu$  ab.

$$\text{Energiedosis } D = \frac{\text{Energie}}{\text{Masse}} \quad (1.1)$$

$$\Psi = \Psi_0 * e^{\mu * d} \quad (1.2)$$

Die Informationen, die nun in der Strahlung vorhanden sind, können durch verschiedene Sensormaterialien registriert und gespeichert werden. Heute wird dazu oft schon digitale Röntgentechnik eingesetzt, bei der, wie bei einer Digitalkamera, die Informationen auf einem CCD Array gespeichert werden. Ein auslaufendes Verfahren ist wohl die konventionelle Röntgenfilmaufnahme, die im Vergleich zur digitalen Aufnahme zwar wesentlich hochauflösender ist, aber sich bei der weiteren Ver- und Bearbeitung als nicht praktikabel erweist und zudem

---

<sup>5</sup>Brustkorb

## 1 Einführung

---

eine hohe Strahlendosis verlangt. Ein Nachteil der Projektionsradiographie ist die Abbildung eines Volumens auf eine Fläche, wodurch Überlagerungen zum Beispiel vom Gewebe des Patienten zustande kommen, die wichtige Bereiche unzugänglich machen können. Hier ist die Computer-Tomographie (CT), ein weiteres radiologisches Verfahren, wesentlich genauer.

### 1.1.2 Computer-Tomographie

Die CT wurde 1967 von dem englischen Physiker Hounsfield entwickelt und wird seit Anfang der 1970er Jahre praktisch angewendet. Es handelt sich hierbei um ein Röntgenschnittverfahren, das durch spezielle Rechenverfahren ein Schnittbild des menschlichen Körpers erzeugt. In einer sogenannten „Gantry“ sind Röntgenröhre und Detektorkranz untergebracht. Die Röntgenröhre rotiert um den Patienten und der Detektorkranz misst die Intensitätsminderung der Strahlen hinter dem Patienten. Hierdurch entsteht ein Volumendatensatz, der durch spezielle Projektion auf einzelne Querschnittsbilder zurückgerechnet wird.

Eine Weiterentwicklung der CT ist die MDR-CT (Multidetektorreihen-Computer-Tomographie). Hierbei werden Reihen von Detektoren verwendet, wodurch mehrere Datensätze gleichzeitig erzeugt werden können und somit die Röntgenstrahlung besser ausgenutzt wird, was eine verminderte Strahlenbelastung für den Patienten zur Folge hat.

### 1.1.3 Magnet-Resonanz-Tomographie und Sonographie

Zu weiteren bildgebenden Verfahren gehören die Magnet-Resonanz-Tomographie (MRT - in der Medizin auch Kernspintomographie genannt) und die Sonographie (Ultraschall). Beide Verfahren werden zur Ergänzung bei der Brustkrebsvorsorge eingesetzt. Die Sonographie kann zum Beispiel echoarme gutartige Zysten von echoreichen

## 1 Einführung

---

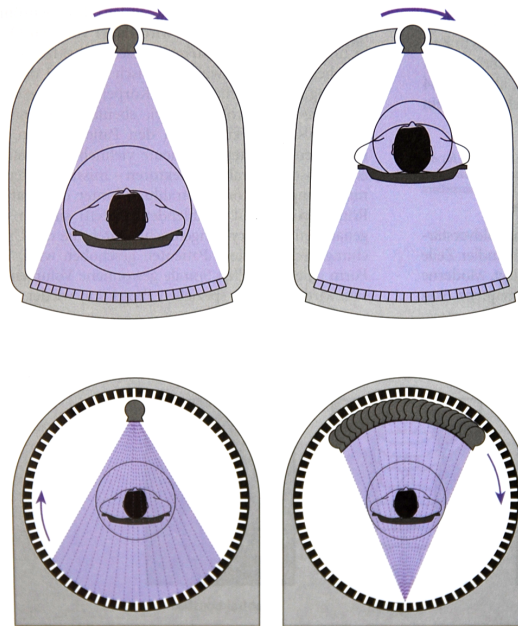


Abbildung 1.1: Computer-Tomographie [8]. Die Röntgenröhre rotiert um den Patienten.

Tumoren unterscheiden. Die MRT wird hingegen oft bei brusterhaltender Therapie zur Differenzierung zwischen Narbe von einer Operation und Rezidiv<sup>6</sup> eingesetzt. Die MR-Mammographie ist ein relativ junges Verfahren, welches sich durch seine hohe Sensitivität aber geringe Spezifität nicht zur Erstuntersuchung eignet.

Bei der MRT werden Atomkerne innerhalb des Körpers an einem Magnetfeld ausgerichtet, und durch ein transversales hochfrequentes Magnetfeld angeregt. Gehen die Atome wieder in ihre Ausgangslage zurück, werden schwache elektromagnetische Felder erzeugt, die ausgelesen werden können und Informationen über den atomaren Aufbau liefern, der dann visuell umgesetzt wird.

Bei der Sonographie oder Ultraschalluntersuchung werden durch einen elektrischen Impuls in der Ultraschallsonde hochfrequente Schallimpulse erzeugt, die im Wellenlängenbereich von ca. 20 kHz bis 1 GHz liegen. Der ausgesendete Schall wird an den inneren Organen entweder absorbiert, gestreut oder reflektiert. Der reflektierte Schall wird

---

<sup>6</sup>Wiederauftreten der Krankheit

## 1 Einführung

---

von der Sonde wieder aufgenommen, je nach Intensität in Grauwerte umgerechnet und zweidimensional als Sonogramm abgebildet. Es gibt mittlerweile Ultraschallverfahren die dreidimensional oder vierdimensional abbilden, also die Zeit noch mit einbeziehen. Beide oben genannten Verfahren haben den Vorteil für den menschlichen Körper nicht schädlich zu sein. Auf beide Verfahren soll in dieser Arbeit nicht weiter eingegangen werden.

Untersuchung	Technische Durchführung	Bedeutung und Grenzen
Mammographie	Röntgenstrahlen	Eingesetzt zur Früherkennung und Abklärung, hohe Evidenz. Problem: „Dichte Brust“
Sonographie	Schallwellen	Eingesetzt zur Abklärung und Früherkennung, hohe Aussagekraft. Problem: Untersucherabhängig
MRT der Brust	Magnetfelder und Radiowellen	Eingesetzt zur Abklärung nach Therapie, bei Hochrisikopatientinnen und lokalem Staging. Höchste Sensitivität. Problem: Spezifität, Untersuchererfahrung.

Tabelle 1.1: Radiologische Verfahren in der Diagnostik.

## 1.2 Die Anatomie der weiblichen Brust

Die weibliche Brust (1.2) besteht im Wesentlichen aus Fett-, Drüsen-, und Bindegewebe, die auf dem großen Brustmuskel (Musculus pectoralis major) liegen. Die Muttermilch wird in den Drüsenlappen (Lobuli glandulae mammae), die wiederum aus mehreren Drüsenläppchen bestehen, gebildet und durch die Milchgänge (Ductus lactifer) zu den Milchsäckchen (Sinus lactiferus) geleitet, die in der Nähe der Mamille<sup>7</sup> als Behälter dienen. Das Drüsengewebe wird von der Haut und der Thoraxwand durch das Fettgewebe getrennt und baut sich mit zunehmenden Alter der Frau fortlaufend ab (Abbildung 1.3).

---

<sup>7</sup>Brustwarze

## 1 Einführung

---

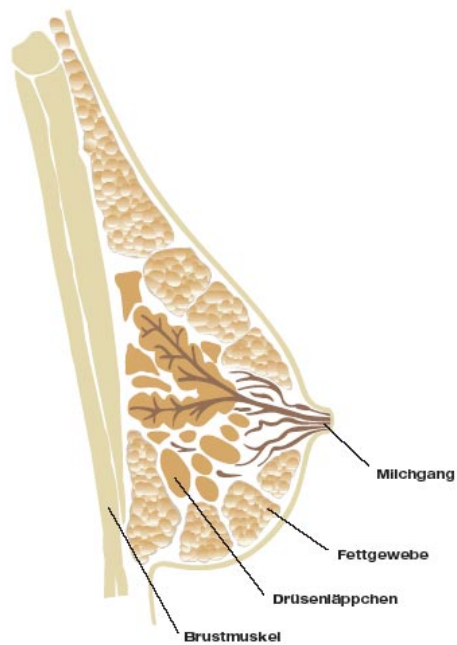


Abbildung 1.2: Der Aufbau der weiblichen Brust.

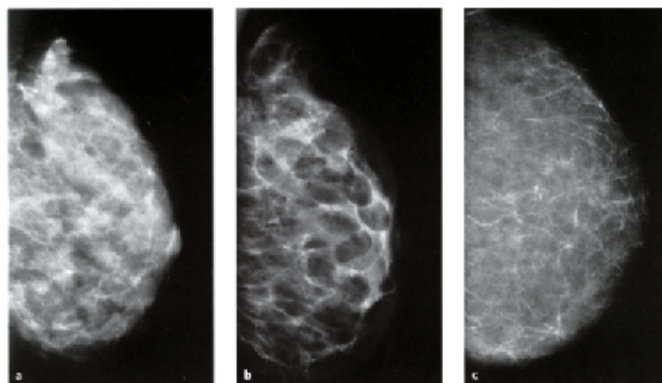


Abbildung 1.3: Zurückbildung des Drüsengewebes mit dem Alter von links nach rechts.

## 1.3 Das Mammakarzinom

Das Mammakarzinom ist ein bösartiger Tumor, der von Zellen im Deckgewebe von Haut oder Schleimhaut ausgeht. Im Allgemeinen wird auch der Begriff Krebs für das Auftreten verwendet. Beim Krebs kommt es zu einem unkontrollierten Zuwachs und Wuchern von Zellen. Normalerweise besitzt der Körper Systeme, die bei Wachstum und Vermehrung von Zellen die Vorgänge kontrollieren und gegebenenfalls korrigieren, also den Zellzyklus überwachen. Diese Prozesse sind aus noch nicht bekannten Gründen außer Kraft gesetzt. Die Zellen vermehren sich also ungehemmt und durch das Aussetzen der Korrekturprozesse kommt es zu immer weiteren Mutationen. Durch diese Mutationen können die Zellen auch die Eigenschaft erreichen „unsterblich“ zu werden oder zu metastasieren<sup>8</sup>. Das Schwierige für das menschliche Immunsystem bei dieser Krankheit ist, dass die mutierten Krebszellen zu sehr den körpereigenen Zellen ähneln, so dass sie nicht wirklich bekämpft werden können. Die meisten Menschen, die an Krebs leiden, sterben an den Folgen der Metastasen, die lebenswichtige Organe befallen, oder generell durch eine Schwächung des Immunsystems aufgrund der hohen Belastung durch den Krebs.

### 1.3.1 Epidemiologie

Das Mammakarzinom ist die am häufigsten vorkommende Krebserkrankung bei Frauen. Im Jahr 2004 starben 17590 Frauen an Brustkrebs, dies ist ein Anteil von 4 % aller Todesfälle bei Frauen und 17 % aller durch Krebs verursachten Tode [1]. Die Inzidenz<sup>9</sup> liegt bei 57000 pro Jahr und die Mortalität<sup>10</sup> je 100000 Frauen in der BRD liegt bei 27 [2]. Die Häufigkeit des Vorkommens von Brustkrebs nimmt mit dem Alter zu, wie auf in Abbildung 1.4 zu sehen ist. Das mittlere Erkrankungsalter liegt bei 63 Jahren.

---

<sup>8</sup>Ausbreiten

<sup>9</sup>Anzahl der Neuerkrankungen

<sup>10</sup>Sterberate

## 1 Einführung

---

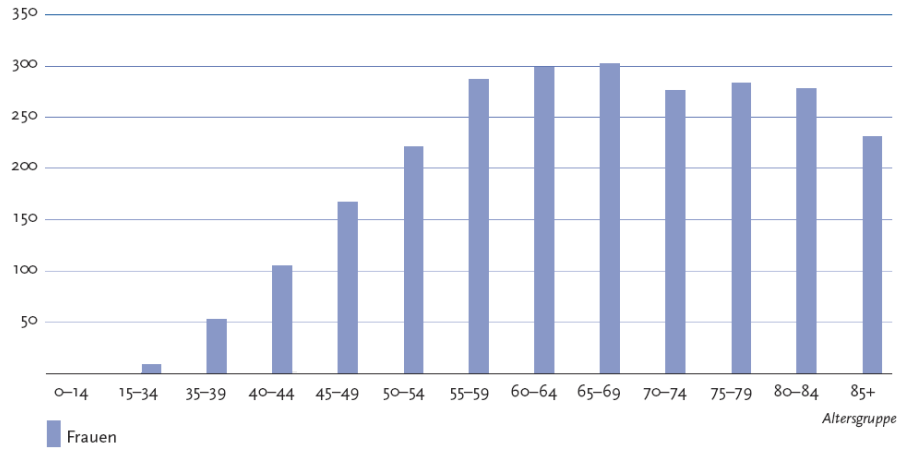


Abbildung 1.4: Schätzung der altersspezifischen Inzidenz der Brustkrebserkrankung in Deutschland 2004, Neuerkrankungen pro 100000 [2].

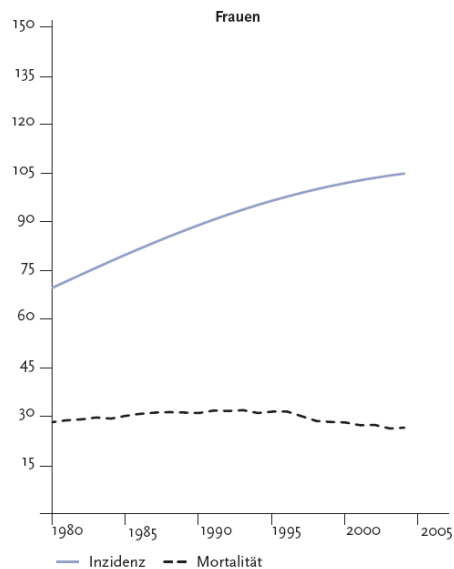


Abbildung 1.5: Inzidenz und Mortalität von Brustkrebs pro 100.000 [2].

### 1.3.2 Symptomatik

Am häufigsten macht die Erkrankung durch schmerzlose Knoten in der Brust auf sich aufmerksam. Es können aber auch schmerzende Knoten, Sekretion oder Einziehung der Mamille, sowie lokale Ödeme<sup>11</sup> auf Brustkrebs hinweisen.

### 1.3.3 Einstufungen des Mammakarzinoms

Das Mammakarzinom wird zuerst grob in invasive und non-invasive Karzinome unterteilt. Etwa 50 % beider Tumorarten werden von Mikrokalzifikationen begleitet. Nur die invasiven Tumore metastasieren, breiten sich also im Körper aus und bilden so genannte Metastasen. Die Ausbreitung in der Brust kann verschiedene Formen annehmen. Es gibt sternförmige Muster, bei denen von einer Verdichtung ausgehend Ausläufer ins umliegende Gewebe wandern, oder knollenförmige Verdichtungen mit unscharfer Begrenzung. Die sich diffus ausbreitenden sternförmigen Tumore sind im Anfangsstadium nur schwer bis gar nicht zu erkennen.

Bei den non-invasiven Tumoren wird zwischen Carcinoma ductale in situ [DCIS] und Carcinoma lobulare in situ [LCIS] unterschieden. Das DCIS bleibt in den Milchgängen und ist ohne begleitende Mikrokalzifikationen bei der Mammographie nicht zu erkennen. Bei ca. 30 % dieser Fälle entsteht ohne Behandlung innerhalb von 5-10 Jahren ein invasives Karzinom [4].

## 1.4 Die Mammographie

Die Mammographie ist die Abbildung der weiblichen Brust durch die schon erwähnte Projektionsradiographie. Die Brust wird dabei an einem speziell angefertigten Röntgengerät fixiert und bis zur Schmerzgrenze komprimiert. Die Komprimierung ist notwendig, um den Dicke-

---

<sup>11</sup>Schwellungen



## 1 Einführung

---

unterschied bei der Aufnahme gering zu halten. Die Mamille wird dabei so angeordnet, dass sie außerhalb der Brust liegt und nicht in das Drüsengewebe hineinprojiziert wird. Es werden grundsätzlich mindestens zwei Aufnahmen von der Brust angefertigt. Eine mediolaterale oblique<sup>12</sup> [MLO] und eine craniocaudale<sup>13</sup> [CC]. Bei der Aufnahme muss noch darauf geachtet werden, dass die Thoraxwand zu sehen ist, um mögliche Folgeaufnahmen vergleichen zu können. In bestimmten Fällen werden auch Vergrößerungen von suspekten Bereichen in der Mammographie angefertigt.

Bei den heutigen Aufnahmesystemen liegt die mittlere Parenchymdosis<sup>14</sup> unter 4 mGy und ist mit keinem bedeutend erhöhtem Risiko einer Krebserkrankung verbunden [4]. Die Röhre ist so angeordnet, dass das Dosismaximum an der Brustwand liegt und bis zur Mamille hin abfällt. Die Belichtung wird heutzutage meist automatisch eingestellt und sollte so bemessen sein, dass auch in einem dichten Drüsenkörper noch Zeichnung vorhanden ist. Strahlungsqualität, Pathologie, Normalbefund so wie weiteres Vorgehen und Konsequenzen aus der Diagnose müssen dokumentiert werden.

Um gute Ergebnisse bei der Diagnose zu bekommen, ist es von Vorteil, wenn Palpation<sup>15</sup>, Mammographie und gegebenenfalls auch Sonographie vom gleichen Radiologen durchgeführt werden. Bei Bedarf kann ein Zweitgutachten von einem weiteren Radiologen in die Diagnose einbezogen werden.

Die beiden wichtigsten Hinweise auf ein Karzinom in einer Mammographie sind sternförmige oder knollenartige Verdichtungen des Gewebes mit spikulären<sup>16</sup> Ausläufern, die in den meisten Fällen von Mikrokalzifikationen begleitet werden. Eine starke Anhäufung von kleineren Mikrokalzifikationen ist ein Zeichen für einen malignen Tumor, aber

---

<sup>12</sup>Schräg vom Kopf nach unten zur Seite hin

<sup>13</sup>Vertikal vom Kopf zum Rumpf

<sup>14</sup>Strahlendosis die vom Gewebe absorbiert wird in Milligray [mGy] gemessen. 1 Gy = 1 J/kg

<sup>15</sup>Tastuntersuchung

<sup>16</sup>Nadelförmigen

## 1 Einführung

---

auch benigne Tumoren treten oft in Begleitung von Mikrokalzifikationen auf, wobei diese oft größer sind und von der Anzahl her geringer.

Weitere Hinweise auf ein Karzinom sind architektonische Störungen des Parenchymusters. Diese macht sich durch zentrale auf einen Punkt bezogene Einziehung des Gewebes bemerkbar [5] und ist nur schwer für den Radiologen zu erkennen. Es kann zum Beispiel ein Tumor unkenntlich unter dichtem Drüsengewebe liegen, wo er jedoch das umliegende Gewebe verzerrt, wodurch ein Verdacht zu Stande kommen kann.

Die Diagnose der Mammographien wird durch einen Radiologen geführt, an den hohe Ansprüche gestellt werden, da der Zeitfaktor begrenzt ist. Studien haben gezeigt, dass sich durch die Zwei-Ebenen-Technik<sup>17</sup> die Sensitivität der Detektion verbessern lässt [21].

Durch die Mammographie können auch Karzinome erkannt werden, die nicht ertastbar sind, wodurch sie sehr wichtig für die Früherkennung der Krankheit ist. Weitere Aufgaben sind die Erkennung der Ausdehnung des Tumors, Unterscheidung zwischen malignen und benignen Tumoren so wie der Rezidiv-Nachweis nach brusterhaltender Karzinomoperation.

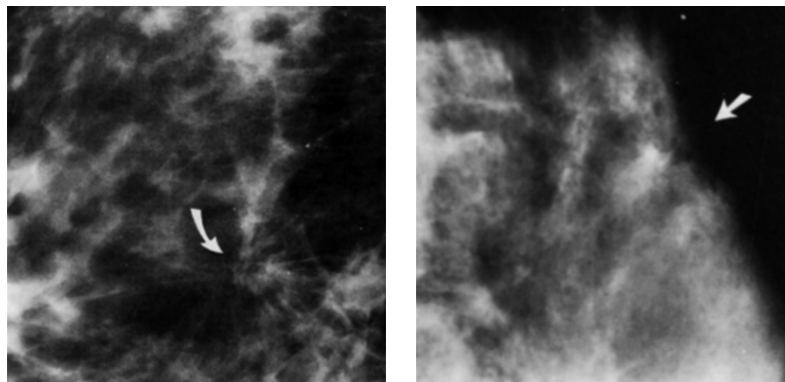


Abbildung 1.6: Architektonische Störungen im Brustparenchym [13]. Das Gewebe ist zu einem Punkt zentriert eingezogen.

---

<sup>17</sup>MLO- und CC-Aufnahme

## 2 Die fraktale Dimension

Dem Menschen sind von Natur aus eher ganzzahlige, die euklidischen Dimensionen ein Begriff. Die fraktale Dimension [FD] ordnet Linien, Flächen und Volumen jedoch, wie der Name schon sagt, gebrochene Dimensionen zu, dass heißt, die Dimension kann hier den Wert rationaler Zahlen annehmen und dadurch zwischen den uns bekannten Dimensionen 1, 2 und 3 liegen. Die einfachste Beschreibung der FD ist die Ähnlichkeitsdimension, die keine Annäherung darstellt, sondern die FD genau ermittelt. Die zur Berechnung notwendigen Parameter sind direkt am Objekt erkennbar.

Man nehme zum Beispiel eine Linie, die man in 5 gleichlange Linien zerlegt. Jede Linie hätte dann  $1/5$  der Länge der ursprünglichen Linie. Macht man das gleiche mit einer quadratischen Fläche, also teilt man das ursprüngliche Quadrat in Quadrate mit einem Fünftel der Seitenlänge auf, so erhält man 25 Kopien des Originals. Bei einem Würfel erhielte man schließlich 125 verkleinerte Kopien des Würfels die selbigen ausfüllen. Wir erkennen also folgenden Zusammenhang:

$$n = \left(\frac{1}{s}\right)^D \quad (2.1)$$

Wobei  $D$  die Dimension,  $s$  der Skalierungsfaktor und  $n$  die Anzahl der Kopien ist. Möchten wir die Dimension ermitteln müssen wir die Gleichung folgendermaßen umstellen:

$$D = \log_s(n) \quad (2.2)$$

$$D = \frac{\log(n)}{\log(s)} \quad (2.3)$$

## 2 Die fraktale Dimension

---

Für eine Linie ergibt sich die Dimension 1, für eine quadratische Fläche die Dimension 2 und für einen Kubus die Dimension 3. Diese allgemeine Gleichung kann auch auf andere selbstähnliche Strukturen übertragen werden, wie zum Beispiel bei dem Sierpinski-Dreieck. Man sieht in Abbildung 2.1 wie drei verkleinerte Kopien des ursprünglichen Dreiecks pro Schritt dazu kommen, die jeweils die halbe Seitenlänge haben. Nach Gleichung 2.2 kommt man auf eine FD von  $D \approx 1,585$ .

Ein weiteres Beispiel ist die Koch-Kurve in Abbildung 2.1, die in jedem Rekursionsschritt vier gleichlange Kopien einer Linie erzeugt. Die Linien entsprechen einem Drittel der Länge der originalen Linie. Daraus ergibt sich eine FD von  $D \approx 0,12619$ . Beide hier aufgeführten Beispiele besitzen die Eigenschaft der Unendlichkeit, da die Rekursionsschritte unendlich oft wiederholt werden können. Die Koch-Kurve wurde schon des Öfteren als mathematisches Monster bezeichnet, da sie trotz ihrer Stetigkeit an keiner Stelle eine Ableitung besitzt.

### 2.1 Das Küstenlinien-Problem

Ein weiteres gutes Beispiel um die FD zu beschreiben ist die Vermessung einer Küstenlinie. 1967 stieß Benoit Mandelbrot, ein Computerwissenschaftler, auf Untersuchungen von L.F. Richardson zur Vermessung von Grenz- und Küstenlinien, die Ausgangspunkt seiner neuen fraktalen Geometrie werden sollten. Die Länge einer Küste ist beim ersten Gedanken die Strecke zwischen zwei Punkten. Eine Küste hat aber eine gebrochene Struktur und die Länge ist somit abhängig davon, wie sehr diese Eigenschaft bei der Vermessung mit einbezogen wird.

Nimmt man beispielsweise das Bild einer beliebigen Insel und beginnt, an einem gewählten Punkt mit einem Zirkel die Küstenlinie entlang zu laufen, wird eine bestimmte Anzahl von Schritten benötigt. Diese Anzahl der Schritte, multipliziert mit dem Radius des Zirkels, ergibt die Küstenlänge. Je kleiner nun der Radius gewählt wird, desto größer wird die Anzahl der Schritte und der gesamten Küstenlänge, da viele

## 2 Die fraktale Dimension

---

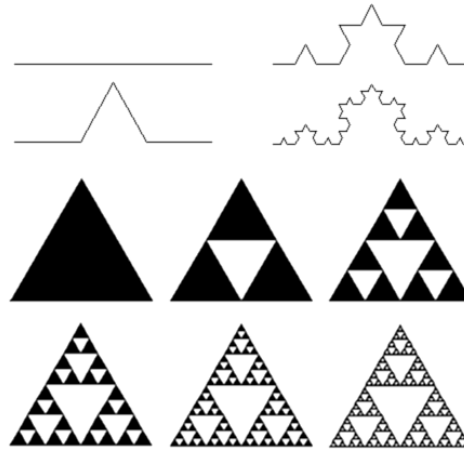


Abbildung 2.1: Iteration der Koch-Kurve (oben) und des Sierpinski-Dreiecks (unten).

kleinere Strukturen nun mit einbezogen werden. Die Küstenlänge ist also abhängig von dem Radius  $\epsilon$  des Zirkels. Wenn man erwartet, dass die Länge der Küstenlinie sich einem bestimmten Wert nähert, so liegt man falsch. Die Länge der Küstenlinie  $L(\epsilon)$  geht für  $\epsilon \rightarrow 0$  gegen unendlich und ist mit Gleichung 2.4 zu bestimmen.

## 2.2 Der Richardson-Effekt

Die Veränderung der Länge der Küstenlinie nach dem erwähnten Verfahren wurden 1961 empirisch von Lewis Fry Richardson untersucht (Abbildung 2.2) [6]. Er fand heraus, dass die Küstenlinie bei einem bestimmten  $\epsilon$  im Wesentlichen von zwei Konstanten abhängt, die im Folgenden als  $\lambda$  und  $D$  bezeichnet werden. Mit diesen Konstanten konnte er näherungsweise die Anzahl der Intervalle  $n$  der Länge  $\epsilon$  feststellen, die notwendig waren, um die Gesamtlänge zu errechnen. Mandelbrot schlug in seiner Arbeit über Küstenlängen vor,  $D$  als Dimension zu interpretieren.

$$n = \lambda \epsilon^{-D} \quad (2.4)$$

Daraus folgt:

$$L(\epsilon) = \lambda \epsilon^{1-D} \quad (2.5)$$

## 2.3 Fraktale Dimension von Oberflächen

Bisher wurden nur Dimensionen zwischen Linien und Flächen in Betracht gezogen. Es kann aber auch von Oberflächen die fraktale Dimension bestimmt werden. Eine gebrochene Fläche liegt zwischen Dimension 2 und 3, und je komplexer diese Oberfläche beschaffen ist, desto höher ist ihre FD. Ein Bild kann im Grunde auch als Oberfläche angesehen werden. Die Pixel bilden die Grundfläche und die Werte der Pixel die Beschaffenheit der Oberfläche, wie in Abbildung 2.3 zu sehen ist. Die FD kann bei diesen komplexen Oberflächen nur näherungsweise bestimmt werden, wie es auch bei der Küstenlinie der Fall ist. Statt der Länge der Küste nimmt hier bei kleinerer Skalierung nun die Oberfläche zu. Die Zusammenhänge bleiben allerdings bestehen, die Gleichung 2.5 muss aber dementsprechend an die Zweidimensionalität angepasst werden (Gleichung 2.6). Ein gutes Beispiel für eine fraktale Oberfläche in der Natur ist die menschliche Lunge. Die gesamte Fläche, an der ein Gasaustausch stattfinden kann beträgt bei mikroskopischer Vergrößerung  $80 \text{ m}^2$ , bei stärkerer Vergrößerung durch ein Elektronenmikroskop  $140 \text{ m}^2$ . Im Grunde kann man sagen, dass die FD die Komplexität der Oberfläche beschreibt.

$$L(\epsilon) = \lambda \epsilon^{2-D} \quad (2.6)$$

## 2 Die fraktale Dimension

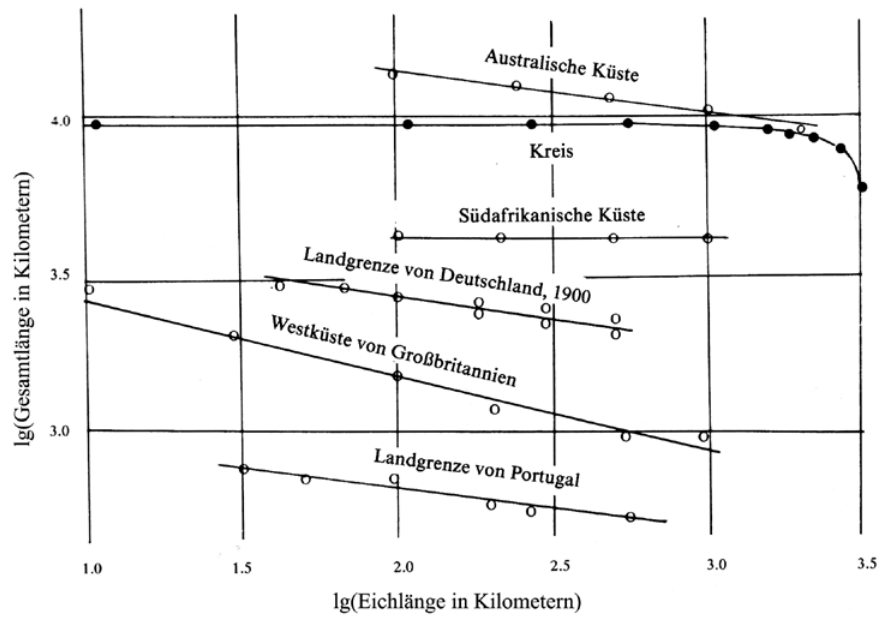


Abbildung 2.2: Richardson-Effekt [6]

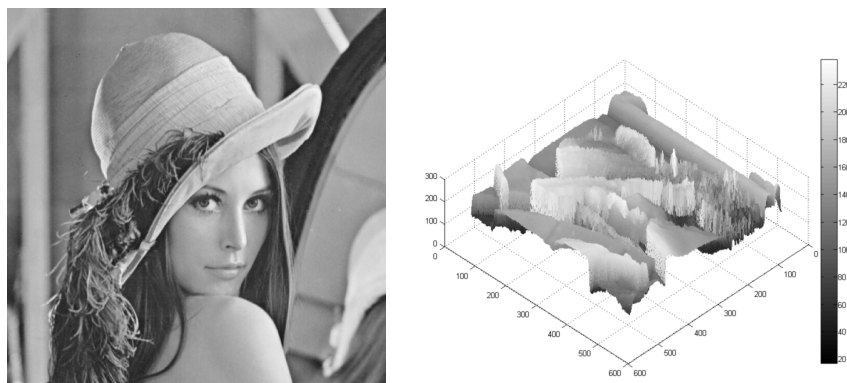


Abbildung 2.3: Lena als zweidimensionales Bild und dreidimensionaler Plot.

## 2.4 Fraktale in der Natur

Fraktale kommen sehr häufig in der Natur vor. Das Wachstum von Schneeflocken, die Ansammlung von Galaxien oder Wolkenbildung sind fraktale Vorgänge, die durch rekursive Algorithmen annähernd beschrieben werden können.

Nimmt man zum Beispiel einen Farn, so sieht man, dass sich an jedem Farnwedel kleinere, in Form identische Kopien des Farnwedels bilden, an denen wiederum kleinere Farnwedel sitzen etc.. Dieser sich wiederholende Prozess folgt einem bestimmten System und ließe sich theoretisch unendlich lang fortführen, ist jedoch in der Natur begrenzt und in dieser Eigenschaft keinem mathematischen Fraktal ähnlich.

Auch bei Wachstumsprozessen von Zellen kommt es zur Bildung von fraktalen Strukturen. Neue Zellen werden gebildet bzw. entstehen bei der Mitose und formen schließlich eine Gewebestruktur, die abhängig von der Funktion oder Art des Gewebes ist. Es liegt nahe, dass sich diese ständige Wiederholung, bei der identische Formen entstehen, in einem selbstähnlichen Muster äußert. Deshalb ist anzunehmen, dass auch bei der weiblichen Brust, die, wie schon erwähnt, überwiegend aus Drüsen- und Fettgewebe besteht, diese Eigenschaft zum Ausdruck kommt und daher auch in der Mammographie zu erkennen sein müsste. Die verschiedenen Gewebearten können unterschiedliche fraktale Dimensionen aufweisen, die sich durch Läsionen im Gewebeaufbau voneinander unterscheiden können, worauf später noch eingegangen wird.

## 2.5 Schätzung der fraktalen Dimension

Es gibt viele verschiedene Algorithmen, um die fraktale Dimension zu schätzen. Da es in dieser Arbeit um diskrete Bilddaten geht, müssen die Verfahren entsprechend der Verwendung auf verschiedene bewährte Methoden eingeschränkt werden, die im Folgenden kurz erläutert werden. Ob diese Methoden jedoch alle geeignet sind aufgrund des



## 2 Die fraktale Dimension

---

Bildmaterials zufriedenstellende Ergebnisse zu liefern sei dahingestellt. Generell liefern die verschiedenen Methoden, auch aufgrund der Tatsache, dass es Schätzungen sind, sehr unterschiedliche Werte. Je nach Bildinhalt reagieren die Algorithmen sehr verschieden und sollten vor Studien gewissenhaft gewählt werden.

### 2.5.1 Methode nach Caldwell

Bei dieser sehr verständlichen Methode nach Caldwell [20] wird das Bild als dreidimensionales Objekt behandelt, wobei der Wert der Pixel die dritte Dimension darstellt. Man kann sich das Bild nun vorstellen wie in Abbildung 2.3 rechts zu sehen ist. Jedes Pixel ist in diesem Modell wie ein Quader anzusehen. Nun wird das Bild nacheinander in quadratische Segmente unterteilt in deren Bereich jeweils der Mittelwert der Pixel gebildet wird. Die Mittelwerte stellen ein neues Bild dar, dessen Oberfläche folgendermaßen berechnet wird. Ist  $\epsilon$  die Seitenlänge eines Segments, und der Wert eines Pixels an der Stelle  $(x, y) = I(x, y)$ , so ist  $A(\epsilon)$  die freistehende Oberfläche der Pixel:

$$\begin{aligned}
 A(\epsilon) = \sum_{x,y} \epsilon^2 + \sum_{x,y} \epsilon [ &|I(x, y) - I(x + 1, y)| \\
 &+ |I(x, y) - I(x, y + 1)|]
 \end{aligned}
 \tag{2.7}$$

Dieser Vorgang wird nun mit bestimmten Segmentgrößen wiederholt. In einer vorhergehenden Studie von Caldwell [8] wurden Segmentgrößen von 4, 9, 16, 25 und 36 Pixel gewählt. Die Größen können entsprechend der Ergebnisse angepasst werden. Im Folgenden wird mit Größen gearbeitet, die einer Potenz von 2 entsprechen. Trägt man nun  $\log(A(\epsilon))$  gegen  $\log(\epsilon)$  auf, so ist die Steigung der Regressiongeraden die fraktale Dimension. Die fraktale Dimension nimmt also mit der Komplexität der Oberfläche zu.

## 2 Die fraktale Dimension

---

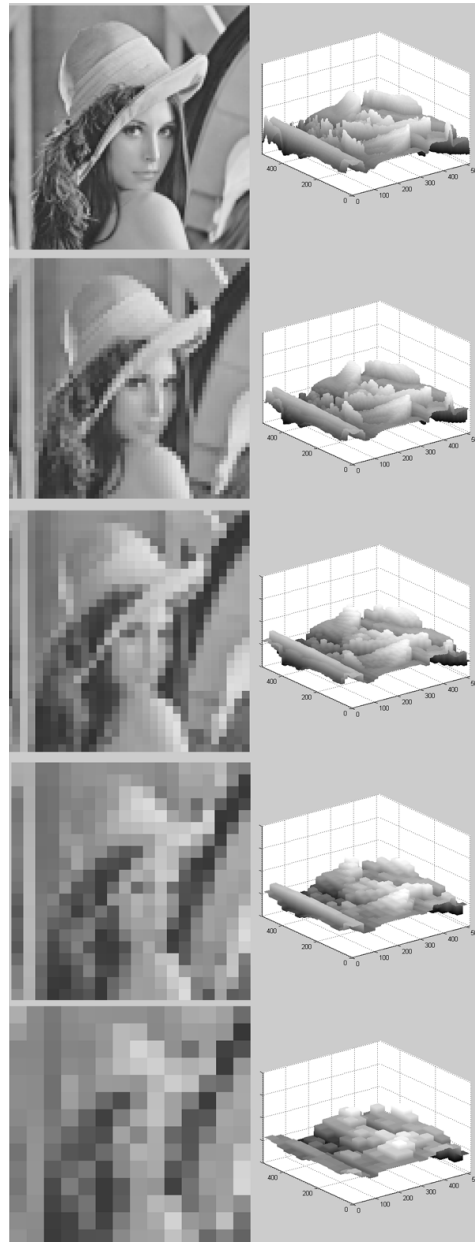


Abbildung 2.4: Caldwell-Algorithmus. Unter verschiedenen Skalierungen wird der Mittelwert gebildet.

### 2.5.2 Circular Average Power Spectrum Method

Bei der Circular Average Power Spectrum Method [CAPS] wird ein quadratischer Ausschnitt aus dem Bild gewählt, dessen Kantenlänge  $2^n$  ist, wobei  $n$  ein Element der natürlichen Zahlen sein muss, da mit der Fast Fourier Transformation [FFT] gearbeitet wird. Bei der FFT wird ein neues Bild mit den gleichen Kantenlängen erzeugt. Das Bild enthält die Amplituden der im Originalbild vorhandenen Frequenzen über verschiedene Richtungen im Bild. Die Werte der Amplituden sind logarithmiert und als Grauwert dargestellt, wie in Abbildung 2.5 zu sehen ist. Diese Grauwerte werden nun vom Mittelpunkt aus für alle Radien des Bildes linear gemittelt und gegen die logarithmierte Ortsfrequenz in einem Diagramm aufgetragen. Man erhält ein eindimensionales Spektrum und verliert die Richtungsinformation. Das eindimensionale Spektrum ist proportional zu  $(\frac{1}{f})^\beta$ , wobei  $f$  die Ortsfrequenz und  $\beta$  die Steigung der Regressionsgeraden ist. Die FD kann nun mit der Gleichung 2.8 [19] berechnet werden. Für die spätere Implementierung hat sich jedoch erwiesen, dass die Gleichung nach Pentland [15] adäquatere Werte liefert. Diese liefert eine um 0,5 niedrigere FD als Gleichung 2.8. Wichtig bei der Berechnung der Steigung ist, dass der annähernd lineare Abschnitt des Diagramms zur Berechnung der Regressionsgeraden herangezogen wird.

$$FD = \frac{7 + \beta}{2} \quad (2.8)$$

Die CAPS Methode ist die wohl am häufigsten verwendete Methode zur Berechnung der FD im medizinischen Bereich. Da es jedoch wichtig ist, das Verhalten der Algorithmen hinsichtlich der zu erhebenden Daten zu kennen, soll später ein Vergleich der Sensitivität verschiedener Algorithmen stattfinden.

## 2 Die fraktale Dimension

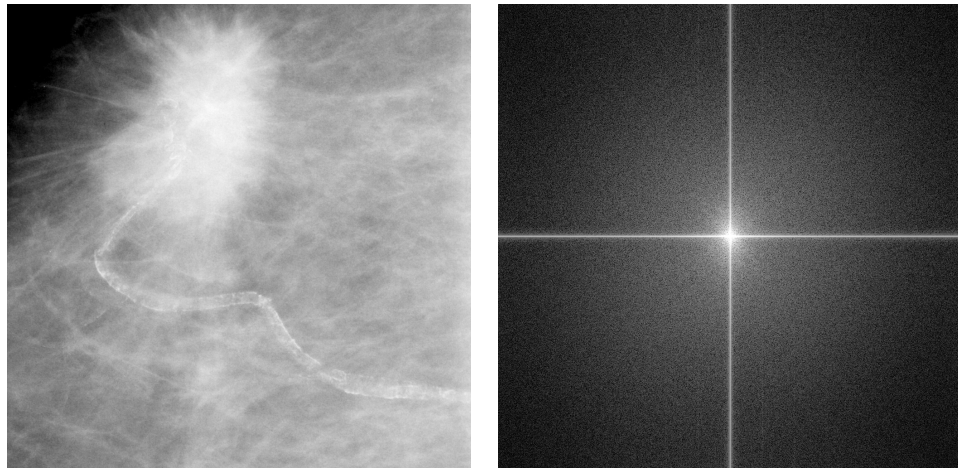


Abbildung 2.5: FFT

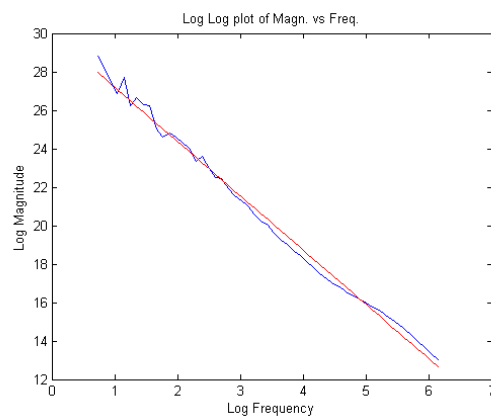


Abbildung 2.6: Bestimmung der FD durch die Steigung der Regressionsgeraden bei der FFT.

### 2.5.3 Triangular Prism Method

Die Triangular Prism Method [TPM] wurde von Clarke (1986) entwickelt, um die Berechnung der FD topographischer Oberflächen durchzuführen. Abseits davon wurde die Methode auch häufig bei der Fernerkundung angewandt, um Oberflächen zu beschreiben. Das Bild wird in ein Raster verschiedener Größe unterteilt, und wie bei dem Caldwell-Algorithmus wird die Oberfläche für verschiedene Rastergrößen berechnet. Jedoch ist die Berechnung der Oberfläche durch Anwendung einer Interpolation genauer. Die Oberfläche der einzelnen Boxen wird berechnet, indem der Mittelwert in der Mitte der Box positioniert wird (Abbildung 2.7). Die Pixelwerte der Pixel an den Ecken der Box ziehen dann mit dem Mittelwert vier Dreiecke auf, deren Fläche addiert die Oberfläche der Box ergibt. Die Oberfläche nimmt mit kleiner werdender Skalierung des Rasters zu, da feinere Strukturen mit einbezogen werden. Trägt man Oberfläche versus Skalierung doppeltlogarithmisch auf, dann entspricht die Steigung der Regressionsgeraden  $2 - D$ , wobei  $D = FD$ .

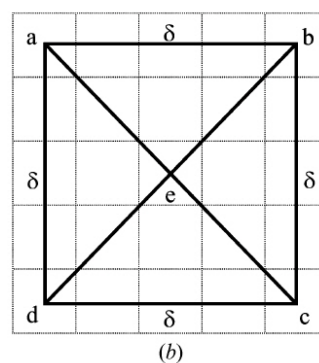
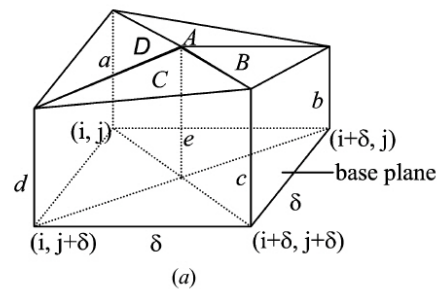


Abbildung 2.7: Triangular prism method

### 2.5.4 Differential Box-Counting Method

Die Differential Box-Counting Method [BC] wurde von Sarkar und Chaudhuri (1992) vorgestellt. Sie funktioniert wie die normale Box-Counting Methode, nur mit dem Unterschied, dass hier der Pixelwert als dritte Dimension mit einbezogen wird. So kann diese Methode, im Gegensatz zur zweidimensionalen Methode, auch auf Bilder mit Grauwerten angewendet werden. Das Eingangsbild muss nicht vorprozessiert und in ein binäres Bild umgewandelt werden. Das macht die BC durchaus interessant für die fraktale Analyse von Mammogrammen. Bei der konventionellen Box-Counting Methode wird ein binäres Bild mit einem Raster überlagert und es werden die Quadranten gezählt, die die Struktur beinhalten. Dies geschieht wieder für mehrere Skalierungen. Je feiner das Raster ist, desto mehr Quadranten überlagern die Struktur des Bildes. Diese Quadranten werden gezählt und gegen die Skalierung doppellogarithmisch aufgetragen. Die Anzahl der Quadranten wird wie bei der Ähnlichkeitsdimension als Anzahl der Wiederholungen bzw. der Kopien behandelt. So wie sich bei dem Sierpinski-Dreieck die Dreiecke in kleinerer Form wiederholen, sind es hier die Boxen oder Quadranten. Die FD ergibt sich direkt aus der Steigung der Regressionsgeraden. Bei der BC ist es so, dass ein dreidimensionales Gitter verwendet wird und für verschiedene Feinheiten dieses Gitters die Kuben gezählt werden, die notwendig sind, um die Struktur zu überlagern. Sei ein Bild  $M * M * M'$  wobei  $M'$  der Wert des Pixels ist und  $M' = M$ , dann wird es mit einem Gitter überlagert, dessen Gitterzellen  $s * s * s'$  groß sind. Jetzt werden die Gitterzellen gezählt, die notwendig sind, um die Struktur zu überlagern. Das heißt, die Gitterzellen müssen die minimalen und maximalen Grauwerte des Ausschnitts umschließen. Die Anzahl in einem Bereich in  $(i, j)$  entspricht  $n_r(i, j) = Z_{max} - Z_{min} + 1$ , wobei  $Z_{min}$  und  $Z_{max}$  die mini- und maximalen Pixelwerte sind. Die Anzahl der Boxen  $n_r$  die dazu nötig sind werden über das ganze Bild aufsummiert (Gleichung 2.9).  $\log(N_r)$  wird gegen  $\log(s/M)$  aufgetragen, wobei die Steigung der Regressionsgeraden  $-D$  entspricht.

$$N_r = \sum n_r(i, j) \quad (2.9)$$

### 2.5.5 Variogram Method

Die Variogram Method [VM] ist eine weitverbreitete Technik zur Bestimmung der FD von Oberflächen. Das Bild wird in ein Raster unterteilt. In den einzelnen Zellen des Rasters wird die Varianz bestimmt. Dieser Vorgang wird für verschiedene Rastergrößen wiederholt. Die Varianz wird doppellogarithmisch gegen die Rastergröße aufgetragen und die FD kann aus der Regressionsgeraden des Diagramms bestimmt werden (Gleichung 2.10).

$$FD = 3 - \frac{b}{2} \quad (2.10)$$

### 2.5.6 Isarithm Method

Die Isarithm Method [IM] basiert auf der Annahme, dass komplexe Oberflächen längere Isolinien (Isarithmen) haben. Isolinien verbinden Punkte gleichen Wertes. Bei einem digitalen Bild wären dies die Pixel mit dem gleichen Grauwert. Nun wird ein Intervall festgelegt, in

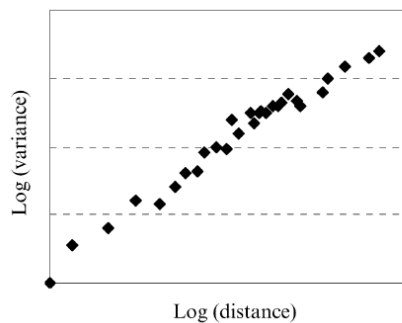


Abbildung 2.8: Variogram Methode. Log(Varianz) vs. Log(distance)

dem die Isolinien berechnet werden. Die Konturlinien, die nun entstehen, werden wie bei der BC Methode analysiert. Alle die Konturlinien überlagernden Boxen werden für verschiedene Skalierungen gezählt und wiederum logarithmisch gegen die Boxgröße aufgetragen. Aus der Steigung der Geraden wird die FD abgeleitet. Schließlich wird die FD der Intervalle, die  $R^2 \geq 0,9$ , gemittelt und 1 addiert.

### 2.5.7 Methodenkritik

Jede Methode hat andere Relationen, die zur Bestimmung der FD herangezogen werden, und je nachdem welche Qualität des Bildes bemessen wird, so verändert sich auch die Schätzung. Tatsache ist, dass in einigen Studien grobe Unterschiede zwischen den Verfahren bei gleichen Voraussetzungen festgestellt wurden (Lam 1990, Clarke und Schweizer 1991, Myint 2003, Sun 2006). Für dieses Zustand kommen sind wohl mehrere Faktoren verantwortlich. Es liegt zum einen daran, dass die einzelnen Methoden verschieden an das Fraktale Modell anknüpfen. Während bei der TPM eine stochastische Methode angewandt wird, ist es bei der IM eine geometrische Herangehensweise. Manche Berechnungen sind sogar nur für selbstähnliche Daten. Bei anderen Bildern ohne selbstähnlichen Inhalt käme es zu groben Fehlern in den Ergebnissen. Bei selbstähnlichen Datensätzen kann es auch vorkommen, dass bei zu grober Skalierung die Dimension falsch angegeben wird. Ist die Größe des Ausgangsobjektes, das sich in kleineren Kopien wiederholt, wesentlich kleiner als die Skalierung, so wird es nicht erfasst.

Aber selbst wenn die adäquate Methode gefunden wurde und der Prozess korrekt ausgeführt wird, kann es immer noch zu groben Unterschieden in den Werten kommen. Dies kommt durch die Parameterwahl zu Stande. Bei der CAPS muss man zum Beispiel beachten, welchen Abschnitt man für die Berechnung der Regressionsgeraden wählt, die Werte können sich durch Einbezug der tiefsten oder höchsten Frequenzen ändern. Der lineare Abschnitt der Funktion liegt meist im mittleren Bereich, kann aber je nach Bildsituation variieren. Bei Methoden, die die Oberfläche mit einbeziehen wie die Caldwell-Methode,



## 2 Die fraktale Dimension

Methode	Benutze Relationen	Formel	Schätzung der FD
TPM	Oberfläche der einzelnen Dächer, Größe der Auswahl	$A(\epsilon) \propto \delta^{2-FD}$	$M = 2 - FD$
DBC	Anzahl der Boxen um Pixelwerte zu überlagern, Boxgröße	$N_r \propto (1/r)^{-\beta}$	$M = -FD$
IM	Länge der Konturlinie, Boxgröße	$L(\epsilon) \propto \delta^{1-FD_{kontur}}$	$M = 1 - FD_{kontur}, FD = \text{Mittelwert von } FD_{kontur} + 1$
Caldwell	Oberfläche, Auswahlgröße	$A(\epsilon) \propto \delta^{2-FD}$	$M = 2 - FD$
VM	Varianz, Distanz	$Var(\delta) \propto \delta^{2-FD}$	$FD = 3 - \frac{M}{2}$
CAPS	Magnitude, Frequenz	$P(f) \propto (\frac{1}{f})^M$	$FD = M + \frac{5}{2}$

Tabelle 2.1: Methoden der Schätzung der FD

## 2 Die fraktale Dimension

---

müssen die Skalierungen entsprechend der Daten gewählt werden, die maximale sowie die minimale Skalierung und die Schrittgröße müssen bestimmt werden. Genaue Vorgaben gibt es dafür nicht, die kleinste Skalierungsgröße liegt aber meist im Bereich der geringstmöglichen, also bei digitalen Bildern der Pixelabstand. Die größte Skalierung wird meist im Bereich  $Bildbreite/2$  bzw.  $Bildhoehe/2$  gewählt.

Weit verbreitet und akzeptiert ist die Größe der Intervalle zwischen den Skalierungen. Diese entsprechen meist einer Zweierpotenz, also einer Verdopplung bzw. Halbierung der Skalierung. Dies ist unter anderem dazu gut geeignet, um bei der doppellogarithmischen Auswertung gleiche Abstände der Daten auf der Abszisse zu haben. Wäre das nicht der Fall und wählte man die Schrittweite linear, so hätte man für grobe Skalierungen weniger Datenpaare als für niedrige.

Eine weitere Frage ist die Anzahl der Schritte. Je mehr Schritte man wählt, desto genauer wird die Messung. Wählt man aber eine Verdopplung der Schrittweite, so steht diese der Anzahl entgegen. Bei einem  $512 \times 512$ px großen Bildausschnitt, könnte man gerade, bei Einbeziehung des geringsten Pixelabstands, neun Schritte prozessieren, wobei die Skalierung auf 512px je nach Ausgangslage nicht notwendig ist, oder auch bei vermuteten selbstähnlichen Strukturen falsche Ergebnisse liefert.

### 2.5.8 Vergleich der Caldwell und CAPS Methode

Um einen geeigneten Vergleich herstellen zu können, wurde ein Programm in Matlab geschrieben, welches von einem bestimmten Bildausschnitt der Kantenlänge  $2^n$  die FD mittels der Algorithmen berechnet und die Ergebnisse in Arrays abspeichert. Die Arrays werden vom Programm in einem Diagramm dargestellt. Es kann die FD der Messung angezeigt werden oder die auf den Wertebereich (0...1) abgebildeten Werte der verschiedenen Methoden (Abbildung 2.10).

Schon beim ersten Versuch mit der CAPS und Caldwell-Methode hat sich gezeigt, dass die Algorithmen sehr verschieden auf das Eingangsbild reagieren und sogar teilweise gegeneinander verlaufen. Zudem haben die geschätzten FD Werte beider Methoden eine hohe Differenz. Als Testbild wurde zuerst ein Ausschnitt einer Mammographie mit spikulären Ausläufern gewählt (Abbildung 2.9). In diesem Beispiel ist gut zu sehen wie unterschiedlich die beiden Verfahren auf die verschiedenen Bereiche im Bild reagieren. Im Bereich der Läsion bei Messung 5, 6 und 7 reagieren beide ähnlich, der Caldwell-Algorithmus scheint nicht so stark auf die Veränderungen anzusprechen wie der CAPS-Algorithmus. Um die beiden Algorithmen noch genauer miteinander zu vergleichen, wurde ein Testbild erstellt, welches acht Stufen hat, die jeweils mit einem gaußschen Weichzeichner verschiedener Stärke bearbeitet wurden. Nun wird von jeder Stufe in einem markierten Bereich die FD nach beiden Methoden berechnet. Hier sieht man deutlich, wie unterschiedlich die beiden Algorithmen reagieren. Während die Caldwell-Methode exponentiell gegen zwei abnimmt, nimmt die FD bei der CAPS Methode mit der Unschärfe im Bild zu. Wenn man sich nochmal die Küstenlinienproblematik als Beispiel zur Berechnung der FD vor Augen führt, widerspricht das eigentlich dem erwarteten Verlauf, da die Komplexität und somit auch die FD der Oberfläche normalerweise abnimmt, je weniger hohe Frequenzen im Bild enthalten sind.

Da die Bestimmung der FD mit der CAPS-Methode über die Regressionsgerade der Magnitude bestimmt wird, könnte es nun sein, dass

## 2 Die fraktale Dimension

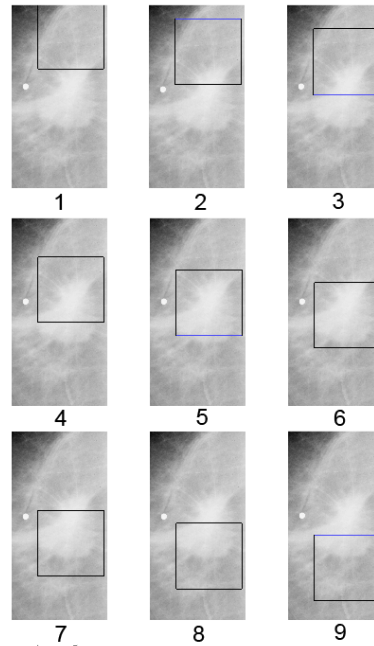


Abbildung 2.9: Messung der FD mit Caldwell- und CAPS-Algorithmus.

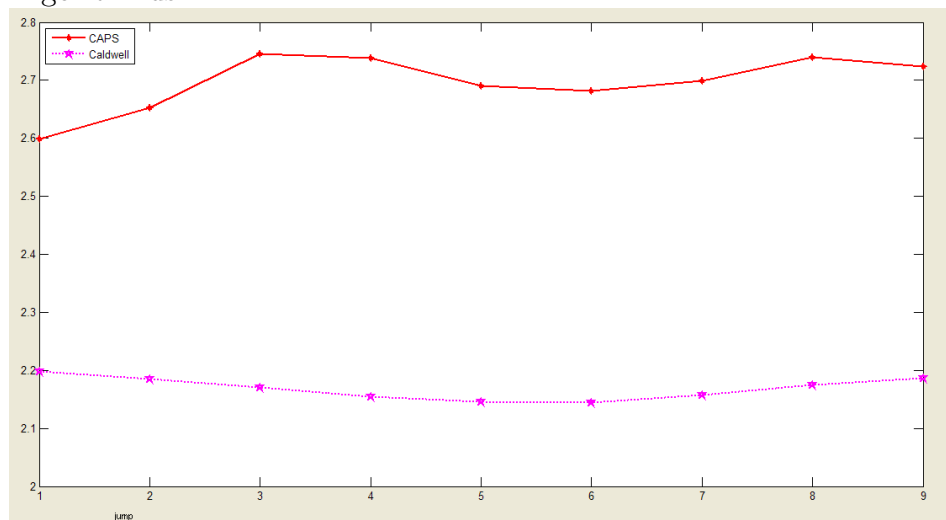


Abbildung 2.10: Die gemessenen Werte aus Abbildung 2.9.

## 2 Die fraktale Dimension

---

wegen der hohen Unschärfe eine Kante in horizontaler und vertikaler Richtung entsteht. Diese Kante würde sehr hohe Frequenzen im Spektrum verursachen und dadurch die errechnete FD vermindern. Um diesen Effekt zu vermeiden wurde vor der FFT eine Fensterfunktion, ein Hanning-Fenster angewendet. Die Fensterfunktion wird mit dem Originalbild multipliziert und sorgt dafür, dass das Ergebnisbild am Rand auf 0 abfällt, wodurch die Kanten verschwinden. Die zweidimensionale Hanning-Fensterfunktion ist in Gleichung 2.11 gezeigt und in Abbildung 2.13 dargestellt. Nun wurde die FD des Testbilds mit dem ungefensterten und dem gefensterten FFT Algorithmus berechnet und auch hier die Werte mit Matlab in einem Diagramm dargestellt (Abbildung 2.14). Man sieht, dass die Fensterfunktion noch eine Verstärkung des Effekts hervorgerufen hat, was im Endeffekt daran liegen muss, dass durch die hohe Unschärfe der Mittelwert des Bildes angenähert wird, der der Frequenz Null entspricht. Diese Frequenz wird in der Berechnung der Regressionsgeraden bzw. der FFT nicht miteinbezogen. Die Fensterfunktion wurde aufgrund dieses Effekts für die weiteren Messungen nicht verwendet.

$$w(u, v) = 0,5 * \cos(\pi * r_{u,v} + 1) \text{ fuer } 0 \leq r_{u,v} \leq 1 \quad (2.11)$$

### 2.5.9 Vergleich weiterer Methoden

Da es wichtig ist die Messungen genau durchzuführen, stellt sich nun die Frage, welcher Algorithmus am besten geeignet ist. Da zwischen Caldwell und CAPS starke Differenzen auftraten, wäre es interessant, noch andere Methoden zu untersuchen. Deswegen wurde noch zusätzlich die BC und TPM Methode implementiert. Für einen Vergleich wurden die Methoden auf ein Testbild mit verschiedenen natürlichen und künstlichen Texturen angewandt. Von den Texturen wurden jeweils die gleichen Ausschnitte von 256\*256px mit den vier Methoden untersucht. Die Messung ist in Abbildung 2.16 zu sehen.

Man sieht deutlich, dass die CAPS Methode stark auf hohe Frequenzen also feine Strukturen im Bild reagiert. Die anderen Verfahren beziehen diese Frequenzen nicht mit ein. Dies liegt an der gewählten Skalierung, die bei dieser Messung bei 4px Boxgröße anfängt. Verringert man die Boxgröße auf 2px, so werden auch feinen Texturen stärker miteinbezogen, und die Ergebnisse der Methoden CAPS, BC und Caldwell nähern sich an. Die späteren Messungen wurden mit einer Boxgröße von 2px durchgeführt.

## 2 Die fraktale Dimension

---



Abbildung 2.11: Testbild mit zunehmender gaußscher Unschärfe

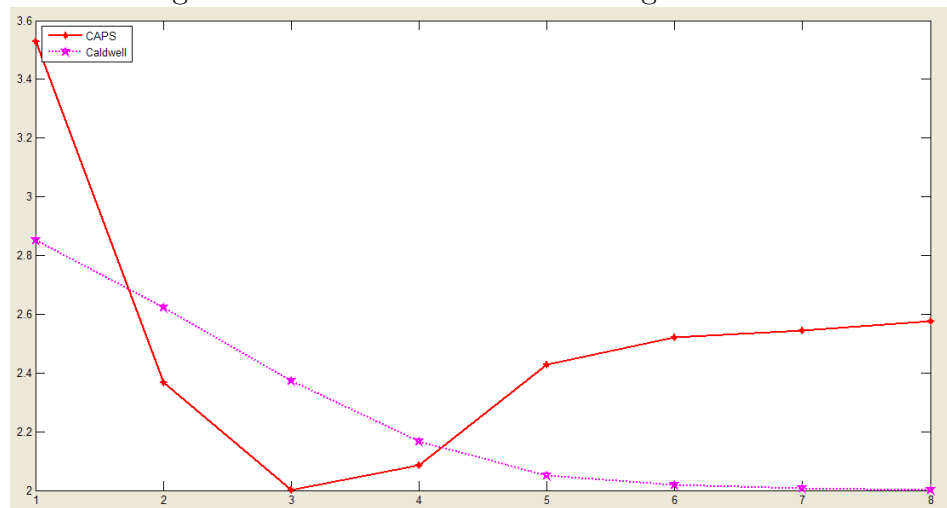


Abbildung 2.12: Die gemessenen Werte aus Abbildung 2.11

## 2 Die fraktale Dimension

---

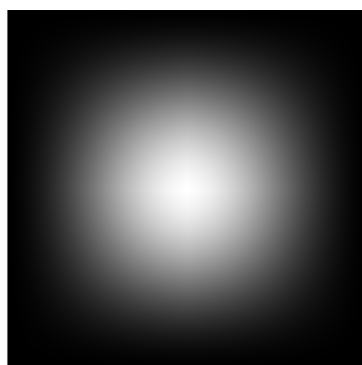


Abbildung 2.13: Hanning-Window Fensterfunktion.

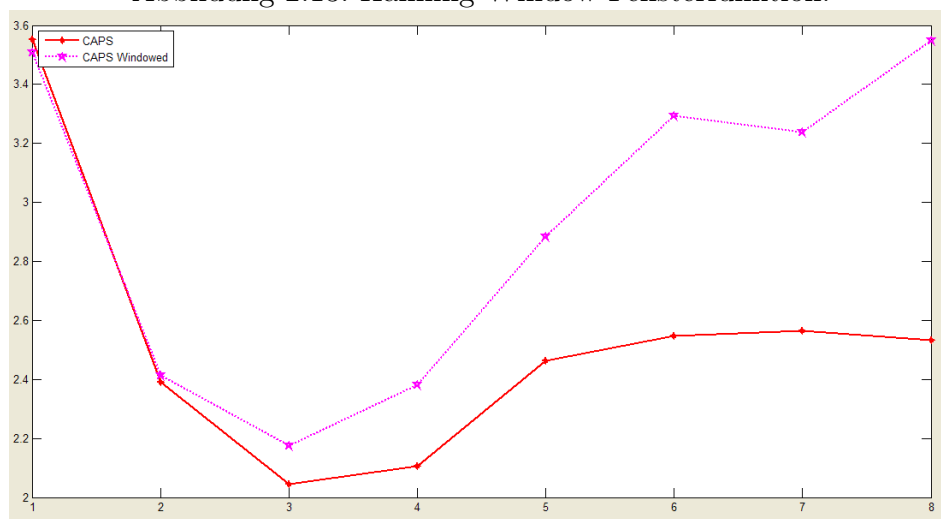


Abbildung 2.14: CAPS-Algorithmus ohne (rot) und mit (magenta) Hanning-Fensterfunktion



## 2 Die fraktale Dimension

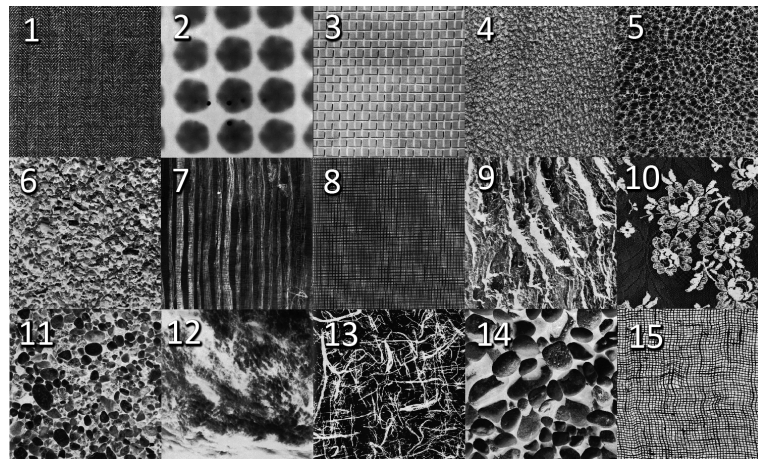


Abbildung 2.15: Testbild verschiedener Texturen

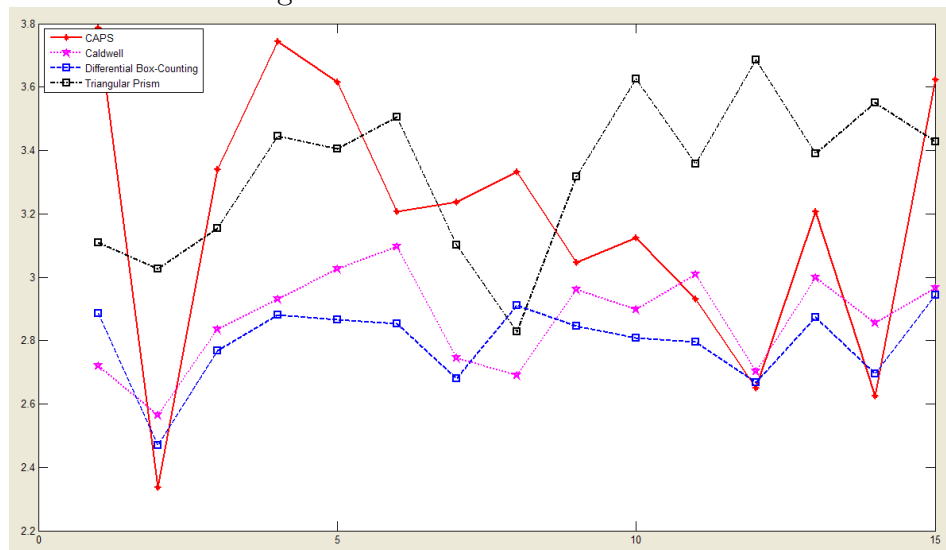


Abbildung 2.16: Vergleich der Messmethoden

## 3 Material und Methoden

### 3.1 Digital Database for Screening Mammography

Die Studie wurde durchgeführt mit digitalisierten Mammogrammen der Digital Database for Screening Mammography [DDSM] [10]. Diese Datenbank der University of South Florida enthält 2620 Fallstudien, die öffentlich zugänglich und speziell für die Entwicklung von Computeralgorithmen zur Auswertung von Mammographien vorgesehen sind. Die Studien sind in 43 Volumen unterteilt, welche normal, benigne oder maligne Fälle enthalten. Jedes Volumen enthält zwischen 20 und 117 Fälle. Die Aufnahmen sind im Lossless Jpeg [Ljpeg] Format abgespeichert und bestehen aus jeweils einer MLO und einer CC Aufnahme der linken und rechten Brust pro Fall. Dazu kommen Informationen zum Patienten, der Zeitpunkt der Studie, die Dichte der Brust<sup>1</sup>, sowie der Befund und dessen Subtilität<sup>2</sup>. Diese Informationen sind in einem .ics File gespeichert. Die .pgm Datei in jedem Ordner dient lediglich der Übersicht und stellt die Mammographien mit Kennzeichnung des Radiologen dar. In vielen Fällen sind auch .overlay Dateien im Ordner enthalten, die diese Kennzeichnungen als numerische Daten enthalten. Für diese Studie war es jedoch praktikabler jeweils einen Screenshot der Thumbnails auf der Homepage der DDSM zu machen. Dort ist die Position der Läsionen ausreichend zu erkennen. Die Screenshots wurden alle mit Befund.png und BefundText.png benannt.

Die Einstufung des Mammographie Befundes folgt dem „Breast image reporting and data system“ [BI-RADS] [9] des American College of

---

<sup>1</sup>density

<sup>2</sup>subtlety

### 3 Material und Methoden

Radiology. Im Englischen steht Assessment für die Einstufung. Die Kategorisierungen werden in Tabelle 3.1 dargestellt.

Kategorie	Bewertung	Beschreibung
1	Mammographie ohne Auffälligkeiten, normaler Befund	Normales Erscheinungsbild
2	Gutartiger Befund	Mammographisch erkennbare Veränderung ohne Hinweis auf das Vorliegen einer malignen Läsion
3	Mit hoher Wahrscheinlichkeit gutartiger Befund	Mit sehr hoher Wahrscheinlichkeit gutartige Veränderung
4	Verdächtige Veränderung, eine gewebliche Untersuchung ist indiziert	Kein charakteristischer, aber möglicher Hinweis auf Malignität
5	Hochgradiger Verdacht der Bösartigkeit, eine Entfernung der Veränderung ist indiziert	Hohe Wahrscheinlichkeit für das Vorliegen eines Karzinoms

Tabelle 3.1: Einstufungen der Befunde nach BI-RADS [18].

Die Subtilität eines Befundes wurde von einem erfahrenen Radiologen nach eigener Einschätzung vorgenommen. Wobei 1 für schwer erkennbar und 5 für offensichtlich steht. Aufnahmen mit Erkennbarkeit 4 und 5 wurden bei dieser Untersuchung außer Acht gelassen.

Des weiteren wird beim Befund zwischen den Erscheinungsformen der Läsion in der Mammographie unterschieden. Es zum einen die Masse selbst, in Form, zum anderen ihr Rand beschrieben (Tabelle 3.2).

Die Dichte der Brust wird wie der Befund nach BI-RADS eingestuft (Tabelle 3.3).

#### 3.1.1 Scanner

Die Mammographien der DDSM stammen von verschiedenen Kliniken und wurden mit drei verschiedenen Scannern erstellt. Jedes dieser

### 3 Material und Methoden

Mass	Margin
Round	Circumscribed
Oval	Microlobulated
Lobulated	Obscured
Irregular	Ill Defined
Architectural Distortion	Spiculated
Tubular	
Lymph Node	
Asymetric Breast Tissue	
Focal Asymetric Density	

Tabelle 3.2: Beschreibung der Läsion anhand von Form der Masse und Qualität des Randes nach BI-RADS [18].

Density	Description
1	The breast is almost entirely fat.
2	There are scattered fibroglandular densities that could obscure a lesion on mammography.
3	The breast is heterogeneously dense. This may lower the sensitivity of mammography.
4	The breast tissue is extremely dense, which lowers the sensitivity of mammography.

Tabelle 3.3: Definition der Brustdichte nach BI-RADS [18].

Systeme hat andere Eigenschaften bezüglich der Aufnahmequalität. Es wurden nur Aufnahmen eines Systems in Betracht gezogen, um gleiche Voraussetzungen bei den Messungen zu haben.

Die Fälle in dieser Studie wurden mit einem Lumisys 200 Laser Scanner an der Wake Forest University [11] eingescannt. Die Scans haben eine Auflösung von  $50 \mu m$  und 12 Bit Farbtiefe. Der Lumisys rechnet die optischen Dichten [OD] linear in Grauwerte um, kann einen Dynamikbereich von 0 bis 3.6 optischen Dichten erfassen und bildet diese auf 3600 Grauwerte [GW] ab. Das nutzt nicht den ganzen 12 Bit Wertebereich. Der Dichtewert 0.0 wird entsprechend auf den Grauwert 4096 abgebildet. Die übrigen optischen Dichten oder Grauwerte ergeben sich nach Gleichung 3.1.

$$OD = 3.6 - \frac{GW - 495}{1000} \quad (3.1)$$

### 3.1.2 Originaldaten

Die ursprünglichen Mammogramme wurden mit einem Kodak Min-R/MRE screen/film System im Department of Radiology an der University of Michigan aufgenommen. Leider gibt es keine weiteren Informationen über dieses System, so dass keine Rückschlüsse auf die tatsächliche Bestrahlung des Films gezogen werden können. Optimal wäre die Kenntnis vom Übertragungsverhalten des Filmmaterials, des Scanners und sonstiger Parameter, die die Informationen beeinflussen, um geräteunabhängige Daten zu gewinnen, die universell anwendbar sind.

### 3.1.3 Konvertierung der Mammographien

Die Mammographien der DDSM sind, wie schon erwähnt, im Ljpeg Format gespeichert. Diese verlustfreie Komprimierung eignet sich nicht zufriedenstellend zur Weiterverarbeitung, weswegen eine Konvertierung ratsam ist. Für die vorliegende Studie wurden die Aufnahmen mit der wxIFSView Software [12] in 8 Bit .tiff Dateien in ursprünglicher Auflösung konvertiert.

## 3.2 Matlab

### 3.2.1 Einleitung

Die Programme zur Auswertung der Messungen wurden in Matlab geschrieben. Matlab ist eine kostenpflichtige kommerzielle Programmierungsumgebung von Mathworks. Sie verfügt bereits über ein sehr umfangreiches Angebot an vorgegebenen Funktionen und Programmen und eignet sich somit gut als Testumgebung. Zudem ist Matlab aufgrund der vorhandenen Image Processing Toolbox und der matrixorientierten Umgebung praktikabel für die Bildverarbeitung. Die Bilder

### 3 Material und Methoden

---

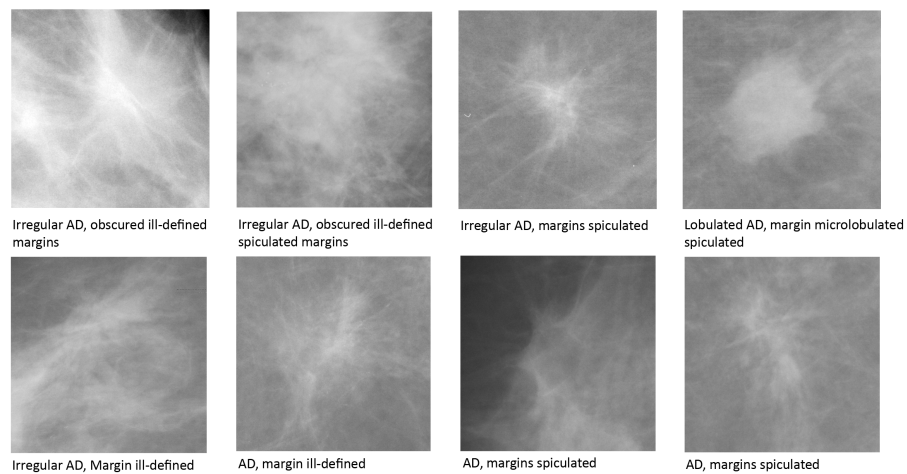


Abbildung 3.1: Beispiele für architektonische Störungen im Brustparenchym. Die Ausschnitte stammen aus der DDSM.

können einfach in den Workspace eingelesen und einer Variablen zugeordnet werden. Jedes Bild wird dabei als Matrix interpretiert. Die vorhandenen Datentypen sind:

- double (0...1.0)
- uint8 (0...255)
- uint16 (0...65535)
- logical (0,1)

Die vorhandenen Bildarten sind

- Index
- Intensity
- Binary
- RGB

### 3.2.2 Funktionen

Die Programme werden in .m-Files gespeichert. Dies sind ASCII-Files, die in einem Editor geschrieben werden. Werden diese Files aus Matlab aufgerufen, so wird der enthaltene Programmcode ausgeführt. Es können, wie auch in anderen Programmiersprachen, Funktionen geschrieben werden. Funktionen können aus einem Programm oder direkt aus dem Commandwindow von Matlab aufgerufen werden. Man kann einer Funktion Argumente übergeben und einen oder mehrere Rückgabeparameter definieren, wodurch es möglich ist, den Programmcode bei umfangreichen Programmen übersichtlich zu halten.

### 3.2.3 Graphical User Interface

Die Programmierung einer grafischen Benutzeroberfläche [GUI] ist in Matlabs GUIDE<sup>3</sup> integriert. Hier können Bedienelemente angeordnet werden, die Funktionen steuern oder auch Eingabeparameter bestimmen. So ist eine sehr einfache und schnelle Bedienung eines Programms gewährleistet. Das angelegte Userinterface wird in einem .fig und einem .m File gespeichert. Der .fig File beinhaltet die grafische Darstellung und Beschreibung der Oberfläche, in dem .m File können in den Callbacks<sup>4</sup> die Bedienelemente mit Funktionen versehen werden. Weitere ausführlichere Informationen zu Matlab gibt es auf der Mathworks Homepage [14].

## 3.3 Programm

### CompareFractalEstimation.m

Das Programm CompareFractalEstimation.m 3.2 vergleicht folgende vier Algorithmen zur Schätzung der FD miteinander, die im Kapitel 2 bereits erläutert wurden.

---

<sup>3</sup>Editor in Matlab zur Erstellung grafischer Oberflächen

<sup>4</sup>Funktion die ausgeführt wird, wenn das Bedienelement betätigt wird

### 3 Material und Methoden

---

- Circular Power Spektrum Methode
- Differential Box-Counting Methode
- Methode nach Caldwell
- Triangular Prism Methode

Ziel des Programms ist es, Aufschluss über die Sensitivität und das Verhalten der einzelnen Algorithmen zu liefern. Dazu wurde eine GUI programmiert, mit der .tif Dateien geöffnet und analysiert werden können. Zur Analyse kann der User eine rechteckige Auswahl, eine Region of Interest [ROI], mit folgenden Seitenlängen bestimmen.

- 128
- 256
- 512
- 1024

Die Auswahl wird mit den Algorithmen analysiert und die geschätzte FD in ein Diagramm gegen die Messung aufgetragen. Hier können nun unterschiedliche Reaktionen an den Messwerten erkannt werden. Zudem besteht die Möglichkeit die Anzeige der Werte der einzelnen Arrays, in denen die Messwerte gespeichert sind, auf den numerischen Bereich (0...1) abzubilden, so dass entgegengesetzte Verhaltensweisen leichter zu erkennen sind.

Um Testbilder mit verschiedenen Eigenschaften zu untersuchen, ist das Programm mit vier Buttons und einem Textfeld ausgestattet, mit denen man die Auswahl um eine bestimmte Pixelanzahl verschieben kann. So kann man Testbilder anfertigen, die zum Beispiel jeweils um 500px versetzt sind. Die ROI kann genau um diesen Wert verschoben werden und trifft somit immer denselben Bildausschnitt. Zudem ist diese Funktion auch nützlich, um die Veränderung der FD zu betrachten, wenn die ROI nur um eine geringe Pixelanzahl verschoben wird, die um ein vielfaches kleiner als ihre Seitenlänge ist.



### 3 Material und Methoden

Mit dem new measurement Button kann man eine neue Messung beginnen. Die Anzahl der Messungen, die über dem Diagramm angegeben ist, wird dabei zurückgesetzt.

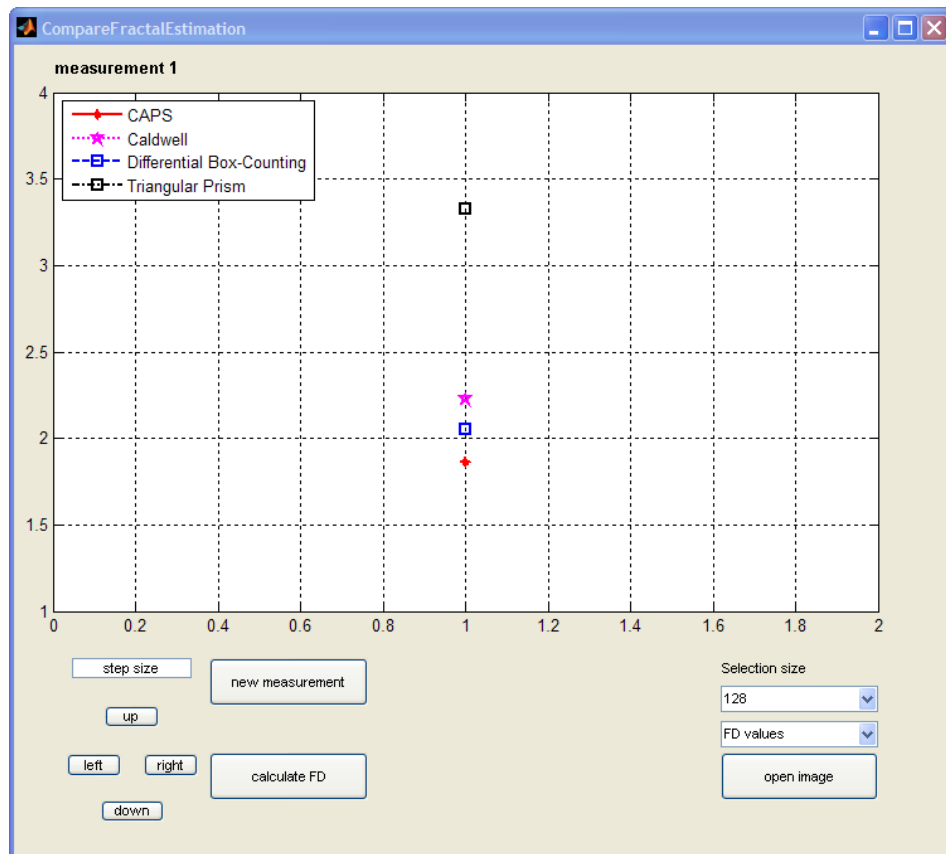


Abbildung 3.2: Programm CompareFractalEstimation.m

## 3.4 Programm

### MammoFractalAnalyzer.m

Das Programm MammoFractalAnalyzer.m dient zur Auswertung der Mammographien aus der DDSM. Es bezieht folgende Algorithmen in die Analyse ein.

- Circular Power Spektrum Methode
- Differential Box-Counting Methode
- Methode nach Caldwell

Die TPM wurde aufgrund stark abweichender Ergebnisse vernachlässigt (Abbildung 2.16).

Das Programm speichert die Messwerte für jede Methode in einem mat-file. Der mat-file wird im Unterverzeichnis Data zur späteren Auswertung abgelegt. Die Messdaten können aber auch als xls-file in diesem Verzeichnis abgespeichert werden.

Das Programm wurde für zwei Monitore ausgelegt. Zuerst öffnet man die Datei Befund.png aus dem Verzeichnis des zu untersuchenden Falles. Das Bild wird auf dem ersten Monitor links unten positioniert. Das Bild BefundText.png wird automatisch in der unteren rechten Ecke angezeigt. Dadurch kann man sehen, wie der Befund lautet und wo sich die pathologischen Veränderungen in der Mammographie befinden. Je nachdem kann nun die gewünschte Aufnahme geöffnet werden. Die Aufnahme wird unten links in einer neuen Figure<sup>5</sup> auf dem ersten Monitor sowie in der GUI oben links dargestellt.

Bevor die Messung begonnen wird, muss ein neuer mat-file mit dem Button new measurement erzeugt werden. Ansonsten erscheint eine Fehlermeldung und das Programm wird nicht ausgeführt. In dem Popmenu Selection size können folgende Größen der ROI festgelegt werden:

- 128

---

<sup>5</sup>graphisches Objekt in Matlab

### 3 Material und Methoden

---

- 256
- 512
- 1024
- 2048

Das Popupmenu Box size legt die Boxgröße fest, in die das Bild unterteilt wird (Abbildung 3.4). Die Auswahlgröße und Boxgröße sollte während der ganzen Messung gleich bleiben.

Die Buttons der GUI haben folgende Funktionen.

„open image“	Öffnet die Aufnahme
„new measurement“	Erzeugt mat-file zur Speicherung der Messdaten
„calculate ad“	Berechnet FD der Läsion
„calculate dg“	Berechnet FD des Drüsengewebes
„calculate fg“	Berechnet FD des Fettgewebes
„save as xls“	Speichert die Messung in einer .xls Datei ab

Tabelle 3.4: Funktionen der Buttons.

Im linken oberen Bild der GUI wird die Auswahl angezeigt und im linken unteren die FD der einzelnen Boxen als Grauwertbild (Abbildung 3.4). Der kleinste FD Wert ist hier auf Schwarz, der höchste auf Weiß abgebildet. Dies soll dazu dienen, erkenntlich zu machen, auf welche Bildsituationen die Algorithmen reagieren, aber auch, ob ein direkter Zusammenhang gesehen werden kann, wenn architektonische Störungen im Gewebe auftreten.

Die Verteilung der FD kann für die verschiedenen Methoden durch ein Popupmenu aufgerufen werden. Zusätzlich wird die mittlere FD der ROI angezeigt.

### 3 Material und Methoden

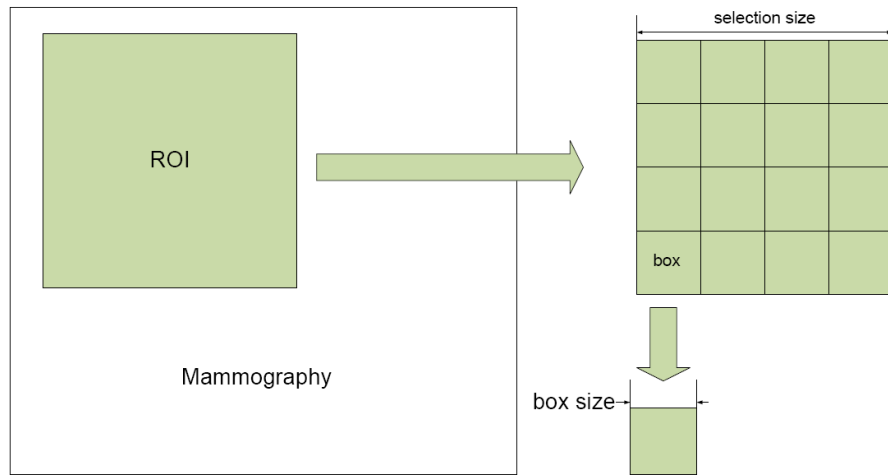


Abbildung 3.3: MammoFractalAnalyzer.m. Die ROI wird in ein neues Bild umgewandelt, welches wiederum in Boxen unterteilt wird.

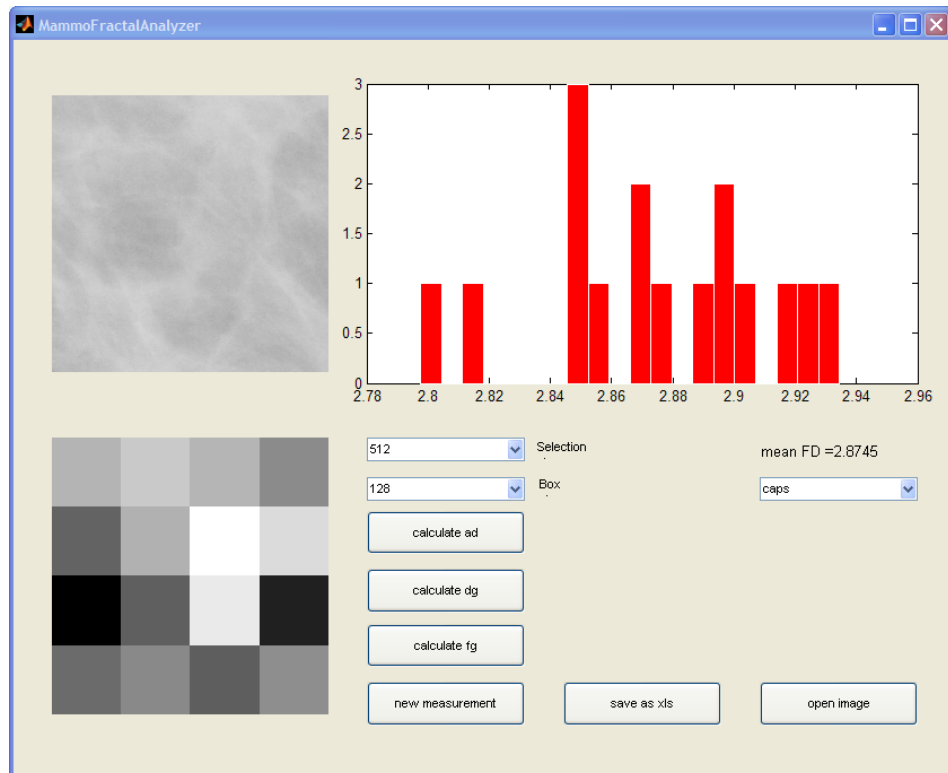


Abbildung 3.4: Programm MammoFractalAnalyzer.m

## 4 Messung und Ergebnisse

### 4.1 Messkriterien

Zuerst einmal ist zu überlegen wie gemessen werden soll. Ziel der Messungen ist es, einen Unterschied zwischen architektonischer Störung und normalem Drüsengewebe bzw. Fettgewebe festzustellen. Also muss jeweils ein Ausschnitt dieser drei Bereiche aus der Mammographie erfolgen. Hierbei besteht das Problem der genauen Lokalisierung der architektonischen Störung, da diese vom Radiologen teils sehr großzügig markiert worden sind. Um die Läsion komplett zu erfassen, reicht in den meisten Fällen eine Auswahlgröße von 512px. Bei einer Auswahlgröße von 256px ist die Markierung des Radiologen in vielen Mammographien schon die ROI überlappend.

Bei kleineren Auswahlgrößen sollte man berücksichtigen, dass auch nur dementsprechend feine Strukturen erfasst werden und nicht die groben Verzerrungen, die eine architektonische Störung ausmachen. Desweiteren strebt die FD für kleinere Auswahlgrößen gegen 2, da sich die Nachbapixel vom Wert her annähern und die „Oberfläche“ des Bildausschnitts glatter wird. Ein weiteres Problem einer kleinen Auswahl besteht darin, dass die Anzahl der Skalierungsschritte geringer wird, was ungenaue Messungen verursachen kann. So ist die Anzahl der Skalierungen bei den benutzten Programmen immer  $\log_2(boxsize)$  was bei einer Boxgröße von 16px nur 4 Schritte bedeuten würde.

In den folgenden Messungen wurde daher mit den Auswahlgrößen 128px, 256px und 512px gearbeitet. Der analysierte Ausschnitt wurde möglichst mittig über der Markierung gewählt. Wurden die Störungen in der MLO und der CC Aufnahme eines Falles als architektonisch gekennzeichnet, so wurden beide Mammographien analysiert.

Der Ausschnitt des Drüsengewebes wurde willkürlich in einem Bereich der Aufnahme gewählt, der möglichst weit von der Störung entfernt lag, die Auswahl aber noch komplett ausfüllte. Die Messung des Fettgewebes erfolgte ähnlich. In manchen Fällen war es jedoch nicht möglich eine Auswahl des Fettgewebes zu untersuchen, da nicht genügend große Bereiche in der Aufnahme vorhanden waren. Diese Messungen wurden dementsprechend vernachlässigt.

## 4.2 Messung

Es wurden 19 Fälle der DDSM untersucht, von denen teilweise mehrere Aufnahmen in die Messungen mit einbezogen wurden. Insgesamt wurden 35 Aufnahmen analysiert. In Tabelle 4.1 ist die Verteilung der Aufnahmen nach Parenchymdichte und Einstufung der Sichtbarkeit zu sehen.

Die Fälle stammen aus den Volumen Cancer01, Cancer02, Cancer05 und Cancer09. Alle Fälle sind nach BI-RADS von 3 bis 5 eingestuft. Für die Untersuchung wurden nur Fälle mit subtlety 3 oder geringer ausgewählt.

Die Aufnahmen wurden mit den bereits vorgestellten Programmen in Matlab untersucht. Es wurde jeweils eine quadratische Auswahl von 256px [Messung A] und 512px [Messung B] des gestörten Gewebes, des normalen Drüsengewebes und des Fettgewebes untersucht.

Qualität	Anzahl	Prozentanteil
Subtlety 1	4	11,43 %
Subtlety 2	16	45,71 %
Subtlety 3	15	42,86 %
Density 1	3	8,57 %
Density 2	7	20,00 %
Density 3	14	40,00 %
Density 4	11	31,43 %

Tabelle 4.1: Verteilung der untersuchten Mammographien.

#### 4 Messung und Ergebnisse

Die FD wurde jeweils mit dem Caps-, Caldwell- und Box-Counting-Algorithmus für die entsprechende Auswahl berechnet.

Eine weitere Messung wurde durchgeführt, bei der eine Auswahl von 512px in 16 Felder mit einer Kantenlänge von 128px unterteilt worden ist [Messung C]. Auch hier wurde die FD für gestörtes, normales und Fettgewebe gemessen.

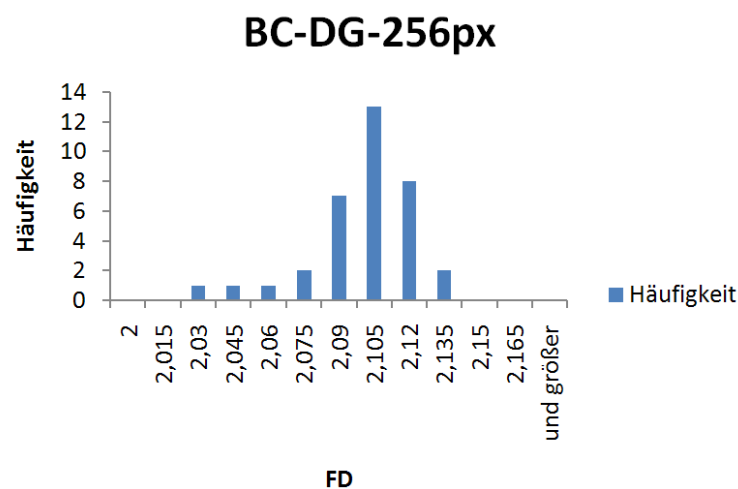


Abbildung 4.1: Verteilung der FD beim Drüsengewebe des Box-Counting-Algorithmus bei Messung C. Man sieht die Tendenz zur Normalverteilung

Die Verteilung der gemessenen Werte wurde nach Darstellung als Histogramm (Abbildung 4.1) als normal verteilt angenommen; Mittelwerte und Standardabweichung der Messungen wurden zur Berechnung und Darstellung einer entsprechenden Normalverteilung ermittelt. Die Normalverteilung für das Drüsengewebe, das architektonisch gestörte Gewebe und das Fettgewebe wurde jeweils in einem Diagramm dargestellt (Abbildung 4.2 - 4.4).

Messung	Beschreibung
A	Auswahlgröße 512px
B	Auswahlgröße 256px
C	Auswahlgröße 512px unterteilt in 16 Boxen mit 128px Kantenlänge

Tabelle 4.2: Messungen

#### 4 Messung und Ergebnisse

---

Die Trefferquote in Tabelle 4.3 wurde nach der Annahme berechnet, dass ein Treffer vorliegt, wenn die FD des gestörten Gewebes niedriger ist, als die des gesunden Gewebes. Bei Messung C wurden die Mittelwerte der FD dazu verwendet.

Summe Treffer	Messung A	Messung B	Messung C
Caps	22	23	21
Caldwell	26	26	19
BC	22	28	14
Trefferquote	Messung A	Messung B	Messung C
Caps	62,86 %	65,71 %	61,8 %
Caldwell	74,29 %	74,29 %	55,9 %
BC	62,86 %	80 %	41,2 %

Tabelle 4.3: Trefferquote.

### 4.3 Ergebnisse

Die Ergebnisse von Messung A und B sind in Tabelle 4.4 und 4.5, die Ergebnisse von Messung C in Tabelle 4.6 zu sehen. Die Normalverteilungen der FD sind in Abbildung 4.2, 4.3 und 4.4 dargestellt.



### 4.3.1 Messung A

CAPS	$\overline{FD}$	$\sigma$	Variationskoeffizient	n
AD 512px	2,67804	0,07329	2,74 %	35
DG 512px	2,71281	0,08201	3,02 %	35
FG 512px	2,88103	0,08045	2,79 %	33
Caldwell				
AD 512px	2,21837	0,03165	1,43 %	35
DG 512px	2,23771	0,02190	0,98 %	35
FG 512px	2,25558	0,02498	1,11 %	33
BC				
AD 512px	2,14412	0,02137	1,00 %	35
DG 512px	2,15815	0,01254	0,58 %	35
FG 512px	2,17200	0,01077	0,50 %	33

Tabelle 4.4: Gemittelte Messdaten von Messung A.  $\sigma$  ist die Standardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.

#### 4 Messung und Ergebnisse

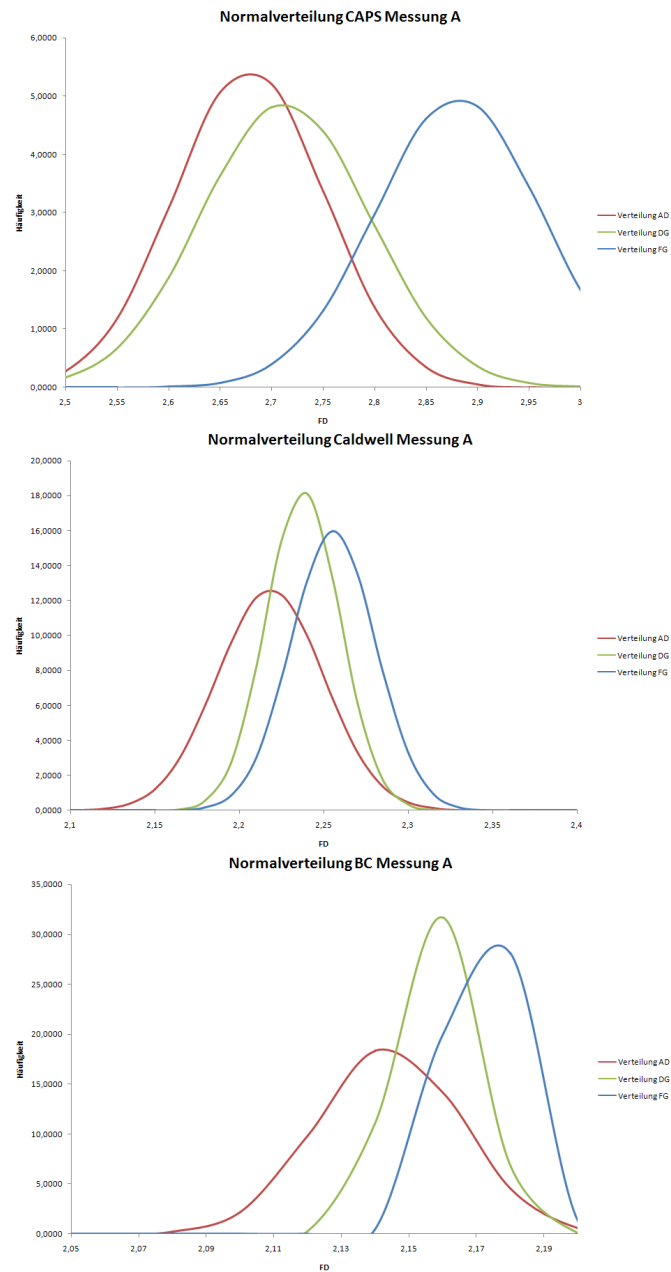


Abbildung 4.2: Normalverteilung Messung A, 512px ROI. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe.

### 4.3.2 Messung B

CAPS	$\overline{FD}$	$\sigma$	Variationskoeffizient	n
AD 256px	2,77278	0,09834	3,55 %	35
DG 256px	2,79030	0,08133	2,91 %	35
FG 256px	2,97336	0,07716	2,59 %	33
Caldwell				
AD 256px	2,16216	0,03075	1,42 %	35
DG 256px	2,18048	0,02128	0,98 %	35
FG 256px	2,19740	0,02375	1,08 %	33
BC				
AD 256px	2,08085	0,03073	1,48 %	35
DG 256px	2,09227	0,02241	1,07 %	35
FG 256px	2,10851	0,01397	0,66 %	33

Tabelle 4.5: Gemittelte Messdaten von Messung B.  $\sigma$  ist die Standardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.

#### 4 Messung und Ergebnisse

---

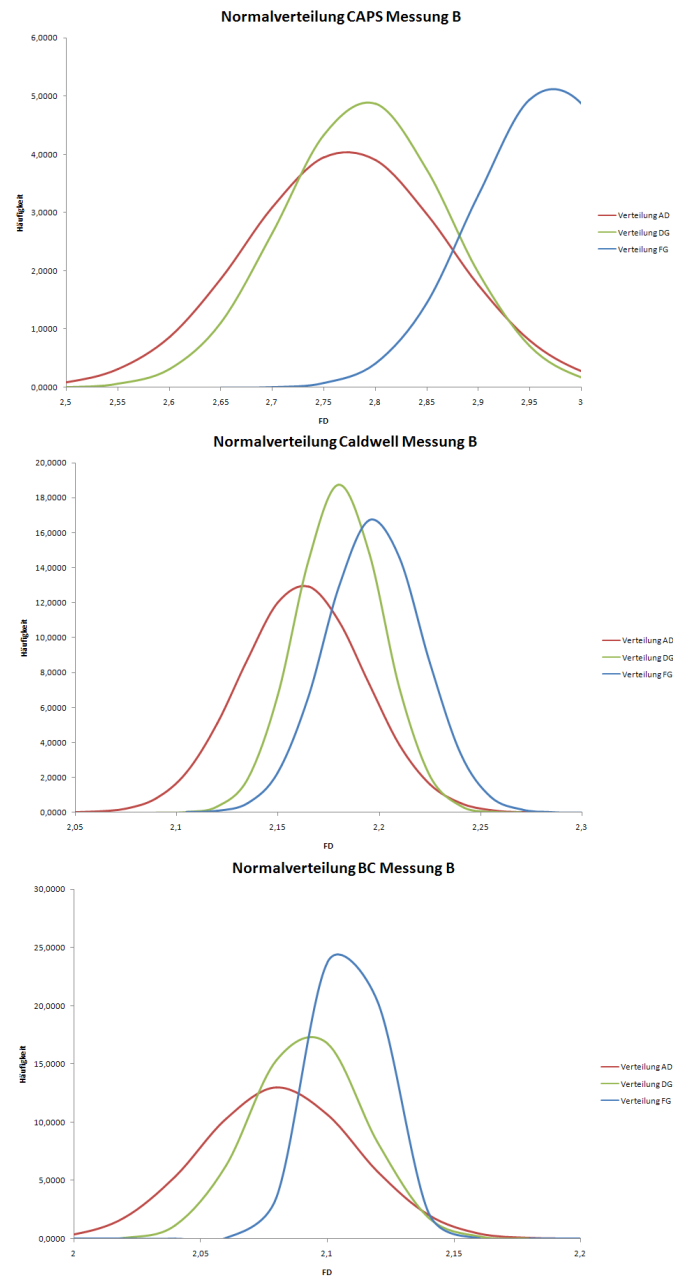


Abbildung 4.3: Normalverteilung Messung B, 256px ROI. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe.

### 4.3.3 Messung C

CAPS	$\overline{FD}$	$\sigma$	Variationskoeffizient	n
AD	2,88095	0,1081	3,75 %	35
DG	2,90376	0,10103	3,48 %	35
FG	3,0436	0,11722	3,85 %	33
Caldwell				
AD	2,11938	0,02360	1,11 %	35
DG	2,12479	0,02190	0,91 %	35
FG	2,1454	0,01686	0,79 %	33
BC				
AD	2,01768	0,13944	0,69 %	35
DG	2,01737	0,01355	0,67 %	35
FG	2,03213	0,01383	0,68 %	33

Tabelle 4.6: Gemittelte Messdaten von Messung C.  $\sigma$  ist die Standardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.

#### 4 Messung und Ergebnisse

---

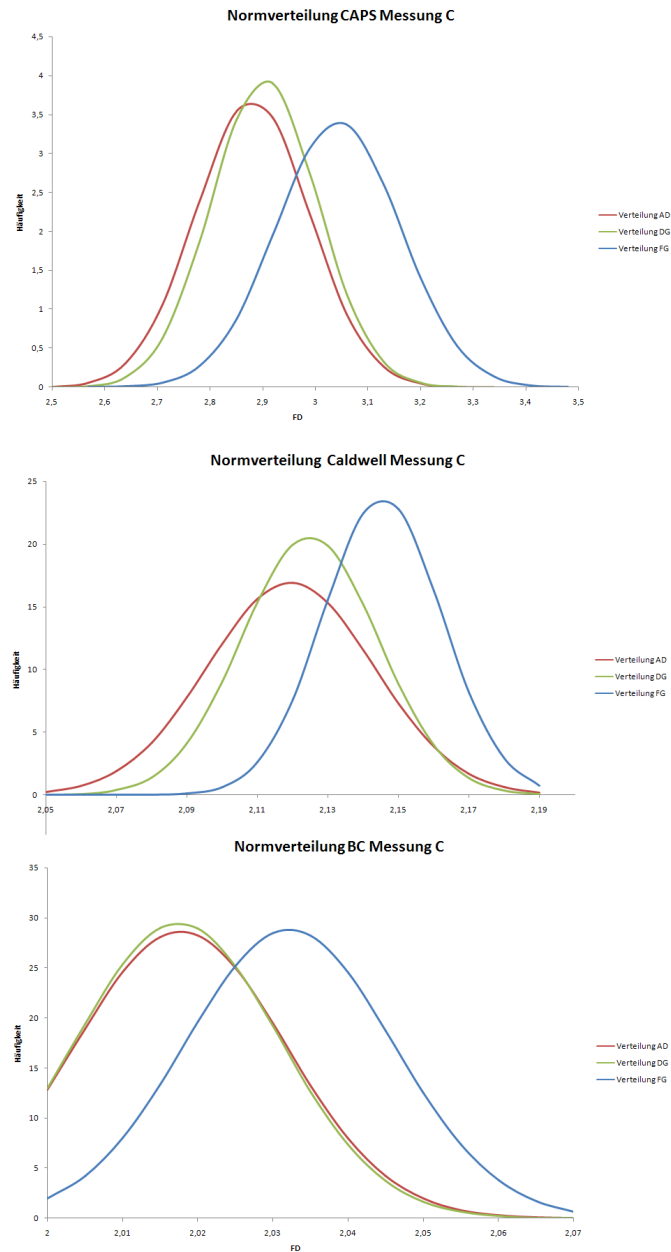


Abbildung 4.4: Normalverteilung Messung C, 512px ROI unterteilt in 128px Boxen. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe.

## 4.4 Statistische Analyse

Wie in den Abbildungen der Normverteilungen ersichtlich ist, kommt es zu starken Überlappungen der Kurven von gestörtem Gewebe und Drüsengewebe. Es ist lediglich eine Tendenz zu erkennen, dass die FD für das gestörte Gewebe im allgemeinen niedriger ist als die des Drüsengewebes, was auch an den Mittelwerten der Stichproben sichtbar wird.

### 4.4.1 t-Test

Um die Differenz quantitativ zu beschreiben, wurde ein t-Test durchgeführt. Bei einem t-Test wird eine Teststatistik unter einer bestimmten Annahme, der Nullhypothese, erstellt. In der Regel ist die Nullhypothese  $H_0$  bei einem Vergleich von zwei Stichproben, dass  $\mu_1 = \mu_2$ , also dass die Erwartungswerte der beiden Proben gleich sind.

Die beiden Stichproben werden als Zufallsvariablen einer Grundgesamtheit angenommen, die einer Normalverteilung entspricht. Nun kann ermittelt werden, wie sehr die Stichprobe von dem Erwartungswert dieser Teststatistik abweicht. Dazu wird aus den Stichproben eine Prüfgröße berechnet. Befindet sich die Prüfgröße unter einem bestimmten vorab festgelegten Signifikanzniveau  $\alpha$ , so trifft  $H_0$  nicht zu. In diesem Fall würde eine statistische Signifikanz vorliegen. Der Ausschlaggebende Wert hierfür ist der p-Wert. Der p-Wert kennzeichnet die Wahrscheinlichkeit bei der angenommenen Teststatistik, dass die Prüfgröße vorkommt oder extremer abweichende Werte. Ist diese Wahrscheinlichkeit kleiner als das Signifikanzniveau, so wird  $H_0$  verworfen und die Alternativhypothese  $H_1 : \mu_1 \neq \mu_2$  angenommen.

Das Ergebnis eines t-Tests stellt keine eindeutige Absicherung einer Nullhypothese dar, sondern beschreibt lediglich, wie hoch die Wahrscheinlichkeit ist, dass  $H_0$  nicht wahr ist.

Für den hier durchgeführten t-Test ist  $H_0 : \mu_1 = \mu_2$ . Zur Berechnung wurde der Zweistichproben t-Test unter der Annahme gleicher Varianzen durchgeführt.

#### 4.4.2 Ergebnisse t-Test

In Tabelle 4.7 sind die Ergebnisse der t-Tests für Messung A, B und C zu sehen. Der t-Test wurde nur zwischen gestörten und normalen Gewebe durchgeführt, da hier eine besonders starke Überlappung bei den Verteilungen zu sehen war.  $P(T \leq t)$  entspricht dem p-Wert.

Messung A	$P(T \leq t)$	Signifikanz bei $\alpha = 0,5$
CAPS	0,066	n.s
Caldwell	0,004	s.
BC	0,0014	s.
Messung B	$P(T \leq t)$	Signifikanz bei $\alpha = 0,5$
CAPS	0,42	n.s
Caldwell	0,005	s.
BC	0,08	s.
Messung C	$P(T \leq t)$	Signifikanz bei $\alpha = 0,5$
CAPS	0,162	n.s
Caldwell	0,257	n.s
BC	0,90	n.s

Tabelle 4.7: Ergebnisse des t-Test, Messung A, B und C.  $P(T \leq t)$  entspricht dem p-Wert.

Aus den Ergebnissen der t-Tests lässt sich ablesen, dass bei einem allgemein üblichen Signifikanzniveau von  $\alpha = 0,05$ , bei Messung A und B, für den Caldwell- und BC-Algorithmus eine statistische Signifikanz vorliegt. Auch sind bei dem CAPS-Algorithmus wesentlich schlechtere Werte zu beobachten, was den größeren Überlappungen der Normalverteilungen in Abbildung 4.2 entspricht. Der Algorithmus hat eine höheren Variationskoeffizient als die anderen Algorithmen, was eine breitere Streuung bedeutet. Die CAPS-Methode scheint generell empfindlicher auf Veränderungen im Bild zu reagieren.



#### 4 Messung und Ergebnisse

---

Die Differenzierung der Gewebearten durch BC und Caldwell wird für kleinere Auswahlgrößen ungenauer und ist für eine Auswahlgröße von 128px, wie in Messung C, nicht mehr statistisch Signifikant. Dies lässt sich vermutlich durch die Annäherung der Pixelwerte in kleineren Auswahlen erklären. Besonders der BC-Algorithmus reagiert stark auf kleinere Bildausschnitte. In Abbildung 4.4 ist gut zu sehen, wie sich die Verteilungen beim BC fast vollständig überlagern. Der CAPS-Algorithmus hingegen verbessert hier zwar leicht die Differenzierung, hat aber trotzdem mit  $p = 0,162$  keine Signifikanz.

Den besten p-Wert erzielte der BC-Algorithmus bei einer Auswahlgröße von 512px mit  $p = 0,0014$ . Auch bei der Trefferquote schnitt er bei dieser Größe mit 80 % am besten ab. Der Caldwell-Algorithmus zeigte bei Messung A und B jeweils eine gute Differenzierbarkeit der Stichproben und liegt mit einer Trefferquote von jeweils 74,3 % in diesen beiden Messungen noch vor der CAPS-Methode mit 62,86 % und 65,71 %.

## 5 Diskussion und Fehleranalyse

Zu den Ergebnissen der t-Tests ist zu sagen, dass hier die Annahme der Normalverteilung beider Stichproben vorausgesetzt wird. Die Wahrscheinlichkeit einer Normalverteilung wurde jedoch nicht untersucht. Es wurde lediglich eine visuelle Auswertung vorgenommen. Speziell bei kleineren Auswahlgrößen, wenn die FD gegen 2 strebt, können andere Verteilungen vorliegen. Auch die Annahme, dass die Stichproben die gleiche Varianz haben, ist nicht bewiesen, aber bei Betrachtung der Standardabweichungen annähernd der Fall.

Ein Fehler in der Schätzung der FD wurde hier nicht behandelt, da es keine Referenz für die Evaluierung der geschätzten FD gab. Eine Referenz könnte zum Beispiel ein Testbild mit bekannter FD darstellen, welches Anhand verschiedener Methoden analysiert wird. Die absolute FD wurde hier deswegen nicht weiter beachtet.

Zu den Daten der DDSM ist anzumerken, dass diese nicht von guter Qualität sind und deshalb auch eine Fehlerquelle darstellen. Artefakte im Bild durch den Scanner oder andere Einflüsse sind keine Seltenheit.

Ein weiteres Problem der Testbilder ist, wie schon in Kapitel 3 erwähnt, dass keine Rückschlüsse auf die ursprüngliche Bestrahlung geschlossen werden können, da die Daten nicht linearisiert sind. Wesentlich genauere Ergebnisse würde man mit digitalen linearisierten Datensätzen bekommen. Wünschenswert wäre hier der Aufbau einer neuen öffentlich zugänglichen Datenbank im Sinne der DDSM.

Zu dem Ergebnis, dass das gestörte Gewebe eine niedrigere FD hat, ist anzumerken, dass auch andere Faktoren als eine Verzerrung im

Gewebe dafür verantwortlich sein könnten. So ist eine architektonische Störung nach DDSM und BI-RADS folgendermaßen definiert:

„The normal architecture is distorted with no definite mass visible. This includes spiculations radiating from a point, and focal retraction or distortion of the edge of the parenchyma. Architectural distortion can also be an associated finding“ [17].

Trotzdem tritt häufig eine leichte Verdichtung bei architektonischer Störung auf. Diese Verdichtung in der Aufnahme hat womöglich auch einen entscheidenden Einfluss auf die FD, da hierdurch im Bild eine eher glatte Oberfläche entsteht. Eine ebenere und somit weniger komplexe Oberfläche hat eine niedrigere FD, was den Ergebnissen in früheren Studien [7] und auch dieser Studie entsprechen würde. Um diese Frage zu beantworten müsste man eine Testreihe zusätzlich mit einem Algorithmus analysieren, der Informationen über die lokale optische Dichte liefert und beobachten, ob hier eine Korrelation zu der FD vorhanden ist.

Zum Abschluß dieser Arbeit ist festzustellen, dass die FD in jedem Fall tendenziell Drüsengewebe von architektonisch gestörten Gewebe unterscheiden kann und dass die Trennschärfe stark beeinflusst von der Wahl des Algorithmus beeinflusst ist. Es ist durchaus denkbar die Methode bei der Computergestützten Auswertung von Mammographien als weiteres Maß mit einzubeziehen. Zu einem besonders guten Ergebnis kam der BC-Algorithmus. Mit einem qualitativ hochwertigeren Datensatz und einer entsprechenden Vorbehandlung der Bilder könnten hier die Ergebnisse durchaus verbessert werden.

Es ist auch vorstellbar diesen Algorithmus auf Volumen-Datensätze einer CT anzuwenden, wobei es hier aufgrund einer wesentlich unterschiedlichen Bildqualität als bei der Projektionsradiographie zu deutlich abweichenden Ergebnissen kommen kann. Dies könnte ein interessanter Ansatz für eine folgende Studie sein.

# Literaturverzeichnis

- [1] Statistisches Bundesamt  
<http://www.destatis.de>  
08/2008
- [2] Krebs in Deutschland  
*Eine gemeinsame Veröffentlichung des Robert Koch-Instituts und der Gesellschaft der epidemiologischen Krebsregister in Deutschland e. V.*  
6. überarbeitete Auflage, 2008
- [3] Pschyrembel - Medizinisches Wörterbuch  
*Willibald Pschyrembel*  
257. Auflage, 1993: S. 933-935.
- [4] Einführung in die Radiologie  
*Egon Bücheler, Klaus-Jürgen Lackner, Manfred Thelen*  
11. überarbeitete und erweiterte Auflage, 2006: S.612-625.
- [5] Breast Imaging  
*Daniel B.Kopans*  
Third edition, 2007: Chapter 13.
- [6] Die fraktale Geometrie der Natur  
*Benoit B. Mandelbrot*  
1.Auflage, 1987.
- [7] A study on the computerized fractal analysis of architectural distortion in screening mammograms  
*Georgia D Tourassi, David M DeLong, Carey E Floyd Jr*  
Published 15th February 2006

- [8] Characterisation of mamographic parenchymal pattern by fractal dimension  
*Curtis B Caldwell et al*  
Phys. Med. Biol., 1990, Vol. 35, No. 2, 235-247
- [9] Radiologie  
*G.W.Kauffmann, E. Moser, R. Sauer*  
3. Auflage, 2006.
- [10] Positive predictive value of the Breast Imaging Reporting and Data System  
*Lacquement, M.A., D. Mitchell, and A.B. Hollingsworth*  
1999: S. 34-40.
- [11] The Digital Database for Screening Mammography  
*Michael Heath, Kevin Bowyer, Daniel Kopans, Richard Moore and W. Philip Kegelmeyer*  
2001
- [12] Wake Forest University  
<http://www1.wfubmc.edu/>  
08/2008
- [13] Department of Electrical and Computer Engineering  
<http://www.ece.ncsu.edu/imaging/Software/index.html>  
08/2008
- [14] Breast Imaging  
*Daniel B. Kopans*  
3. Auflage, 2007.
- [15] The MathWorks  
<http://www.mathworks.com>  
08/2008
- [16] Fractal-based descriptions of natural scenes  
*Pentland, A.P.*  
1984.

- [17] Single reading with computer-aided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program.  
*Gilbert FJ et al.*  
Radiology 2006
- [18] DDSM: Terminology  
[http://marathon.csee.usf.edu/Mammography/DDSM/ddsm\\_terminology.html](http://marathon.csee.usf.edu/Mammography/DDSM/ddsm_terminology.html)  
08/2008
- [19] Breast Imaging Reporting and Data System  
Copyright 1994-2001 American College of Radiology  
2001
- [20] Fractal analysis of remotely sensed Images: A review of Methods and applications  
*W.SUN, G.XU, P.Gong, S.Liang*  
11/2006
- [21] Characterisation of mammographic parenchymal pattern by fractal dimension  
*Curtis B Caldwell, Sandra J Stapleton, David W Holdsworth, Roberta A Jongt, William J WeiserS, Gabriel CookeS and Martin J Yaffe*  
1989
- [22] Observer variability in cancer detection during routine repeat (incident) mammographic screening in a study of two versus one view mammography.  
*Blanks, R.G., M.G. Wallis, and R.M. Given Wilson*  
1999

# Abbildungsverzeichnis

1.1	Computer-Tomographie [8]. Die Röntgenröhre rotiert um den Patienten. . . . .	10
1.2	Der Aufbau der weiblichen Brust. . . . .	12
1.3	Zurückbildung des Drüsengewebes mit dem Alter von links nach rechts. . . . .	12
1.4	Schätzung der altersspezifischen Inzidenz der Brustkrebskrankung in Deutschland 2004, Neuerkrankungen pro 100000 [2]. . . . .	14
1.5	Inzidenz und Mortalität von Brustkrebs pro 100.000 [2].	14
1.6	Architektonische Störungen im Brustparenchym [13]. Das Gewebe ist zu einem Punkt zentriert eingezogen. .	17
2.1	Iteration der Koch-Kurve (oben) und des Sierpinski-Dreiecks (unten). . . . .	21
2.2	Richardson-Effekt [6] . . . . .	22
2.3	Lena als zweidimensionales Bild und dreidimensionaler Plot. . . . .	22
2.4	Caldwell-Algorithmus. Unter verschiedenen Skalierungen wird der Mittelwert gebildet. . . . .	26
2.5	FFT . . . . .	27
2.6	Bestimmung der FD durch die Steigung der Regressionsgeraden bei der FFT. . . . .	27
2.7	Triangular prism method . . . . .	28
2.8	Variogram Methode. Log(Varianz) vs. Log(distance) . .	30
2.9	Messung der FD mit Caldwell- und CAPS-Algorithmus.	35
2.10	Die gemessenen Werte aus Abbildung 2.9. . . . .	35
2.11	Testbild mit zunehmender gaußscher Unschärfe . . . .	37
2.12	Die gemessenen Werte aus Abbildung 2.11 . . . . .	37

2.13	Hanning-Window Fensterfunktion. . . . .	38
2.14	CAPS-Algorithmus ohne (rot) und mit (magenta) Hanning- Fensterfunktion . . . . .	38
2.15	Testbild verschiedener Texturen . . . . .	39
2.16	Vergleich der Messmethoden . . . . .	39
3.1	Beispiele für architektonische Störungen im Brustpa- renchym. Die Ausschnitte stammen aus der DDSM. . .	46
3.2	Programm CompareFractalEstimation.m . . . . .	49
3.3	MammoFractalAnalyser.m. Die ROI wird in ein neues Bild umgewandelt, welches wiederum in Boxen unter- teilt wird. . . . .	52
3.4	Programm MammoFractalAnalyzer.m . . . . .	52
4.1	Verteilung der FD beim Drüsengewebe des Box-Counting- Algorithmus bei Messung C. Man sieht die Tendenz zur Normalverteilung . . . . .	55
4.2	Normalverteilung Messung A, 512px ROI. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe. . . . .	57
4.3	Normalverteilung Messung B, 256px ROI. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe. . . . .	59
4.4	Normalverteilung Messung C, 512px ROI unterteilt in 128px Boxen. AD (rot) ist das gestörte Gewebe, DG (grün) das Drüsengewebe und FG (blau) das Fettgewebe. .	61



# Tabellenverzeichnis

1.1	Radiologische Verfahren in der Diagnostik. . . . .	11
2.1	Methoden der Schätzung der FD . . . . .	31
3.1	Einstufungen der Befunde nach BI-RADS [18]. . . . .	43
3.2	Beschreibung der Läsion anhand von Form der Masse und Qualität des Randes nach BI-RADS [18]. . . . .	43
3.3	Definition der Brustdichte nach BI-RADS [18]. . . . .	44
3.4	Funktionen der Buttons. . . . .	51
4.1	Verteilung der untersuchten Mammographien. . . . .	54
4.2	Messungen . . . . .	55
4.3	Trefferquote. . . . .	56
4.4	Gemittelte Messdaten von Messung A. $\sigma$ ist die Stan- dardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.	56
4.5	Gemittelte Messdaten von Messung B. $\sigma$ ist die Stan- dardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.	58
4.6	Gemittelte Messdaten von Messung C. $\sigma$ ist die Stan- dardabweichung, der Variationskoeffizient die relative Standardabweichung und n die Anzahl der Messungen.	60
4.7	Ergebnisse des t-Test, Messung A, B und C. $P(T \leq t)$ entspricht dem p-Wert. . . . .	63

## Eidesstattliche Erklärung

Ich versichere hiermit, dass ich meine Diplomarbeit mit dem Thema

*„Fraktale Analyse von Mammographien“*

selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Die Ergebnisse der Arbeit unterliegen keinen Sperrvermerk.

Mir ist bekannt, dass ich meine Diplomarbeit zusammen mit dieser Erklärung fristgemäß nach Vergabe des Themas in dreifacher Ausfertigung und gebunden im Prüfungsamt der FH Köln abzugeben oder spätestens mit dem Poststempel des Tages, an dem die Frist abläuft, zu senden habe.

Köln, den 28. August 2008 \_\_\_\_\_

Kai Elbers